Maximum Likelihood Estimation

of Polyserial Correlation

With

Elliptical Variates

By

Wing-Lit Wong

A

Thesis

Submitted to

(Division of Statistics)

The Graduate School

of

The Chinese University of Hong Kong

In Partial Fulfilment

of the Requirements for the Degree of

Master of Philosophy

(M. Phil.)

May, 1989

THE CHINESE UNIVERSITY OF HONG KONG

GRADUATE SCHOOL

The undersigned certify that we have read a thesis, entitled "Maximum Likelihood Estimation of Polyserial Correlation with Elliptical Variates" submitted to the Graduate School by Wing-Lit Wong (黄榮烈) in partial fulfilment of the requirement for the degree of Master of Philosophy in Statistics. We recommend that it be accepted.

Dr. S.Y. Lee,
Supervisor

Dr. N. N. Chan

Dr. W. Y. Poon

Prof. P. C. Chang
External Examiner

DECLARATION

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institution of learning.

ACKNOWLEDGEMENT

The author wishes to thank many people for their enthusiasm and enormous help, including my supervisor, Dr. Sik-Yum Lee.

To all staff of the Department of Statistics, especially two assistant computer officiers of computer laboratory, I extend my sincere thanks. They are always eager to provide assistance whenever I need help most.

Finally, I have to thank my family for their everlasting support.

ABSTRACT

This thesis investigates the maximum likelihood estimation of the parameters in a bivariate polyserial correlation model. The method of finding the maximum likelihood estimates of the parameters in this model, with a component variable obtained in polytomous form and a variable observable in continuous form, is developed. The underlying random continuous variables are assumed to have a bivariate elliptical distribution. Situations that based on two members of the elliptical family, namely, bivariate t distribution and bivariate contaminated normal distribution are studied in details. The Fletcher-Powell algorithm is implemented to find the maximum likelihood estimates. As bivariate normality is a usual assumption in the literature, the robustness of the estimates against the normality assumption is interesting to be investigated. Therefore, finally, simulation studies are conducted for investigating the robustness of these estimators against the normality assumption.

CONTENTS

Chapter 1 Introduction

In many behavioral and social science studies, investigators frequently come across with measurements that only coded in dichotomous or polytomous form. Examples of measurements are performance items, attitude items and satisfaction items. Typical illustration is the Likert scale reported on public attitude towards a controversial statement: (1) Strongly agree (2) Agree (3) Neutral (4) Disagree (5) Strongly disagree. Nevertheless, it is assumed that the observed dichotomous or polytomous data are the results of measurements of some underlying continuous random variables with thresholds which divide the latent continuous variables into ordinal data. But other measurements are obtained in continuous form such as height and time period items. For a deeply study and analysis, it is often neccessary and interesting to exploit such model in more details. One of the important parameters to be investigated is the polyserial correlation, it is the correlation between X and Y obtained from observed X and Z, where Z is an observed discrete variable which depends on the value of an underlying latent continuous random variable Y and thresholds, and X represents another observed continuous variable. Consequently, the parameters in this model will include the polyserial correlation between X and Y and the thresholds.

Under the normality assumption on the distributions of the latent variables and when Z is dichotomous, the maximum likelihood estimation of the correlation between X and Y has been studied by Tate(1955). Later, Tates' work has been generalized by Hannan and Tate(1965). In their researches, they obtained the maximum likelihood estimates of the correlation, thresholds and the standard error estimates. By treating Z as a polytomous observed variable, Cox(1974) generalized Tates' work further and obtained maximum likelihood estimates of the parameters via the scoring algorithm. Olsson, Drasgow and Dorans(1982) compared the maximum likelihood estimator with a two-step estimator and with a simple ad hoc estimator. In a more generalized context and under the normality assumption, Lee and Poon(1986) developed a method for estimating the parameters of the polyserial correlation model in which X is an observable random vector and Z is an observable ordinal polytomous variable. The parameters estimated in the model contain the mean vector, covariance matrix of X, the thresholds and the polyserial correlations between X and Y. Recently, Poon and Lee(1987) generalized the model to the situation where Z is an observable polytomous random vector, based on values of Y and the thresholds.

However, as the results cited above depend much on the assumption of normality, validity of this assumption is also fatal to the validity of the results. Therefore, it should be checked before analyzing the data. Moreover, we do not know how bias are the results when the normality assumption is violated. Will the behaviour of the estimates vary a lot if the assumption is violated? Thus, it is desirable to study the estimation of the polyserial correlation and thresholds based on some other more generalized distributions. One of the most important generalization is to extend the theories to the class of elliptical distribution, as it not only includes the bivariate normal distribution, but also contains platkurtic and leptokurtic distributions. This means that the distributional assumption is much less restrictive.

In this thesis, the polyserial correlation model is described, the maximum likelihood method and the theory for estimating the polyserial correlation and thresholds under the assumption of two members of elliptical distributions are studied in chapter 2. The optimization procedure is included in this chapter. In chapter 3, the structure of the program and technical details are given. Methods of generating data and outline of a simulation study are presented in chapter 4. Finally, as one aim of simulation studies is to investigate the robustness of the normality assumption, the discussion of the results and the conclusion are given in chapter 5.

Chapter 2. The polyserial correlation model

§2.1 General theory of ML estimation on elliptical family

. Let $\underline{T}=(X,Y)'$ be a continuous random vector with mean vector $E(\underline{T})=\underline{0}$, and $\mathrm{var}(X) = \mathrm{var}(Y) = 1$, $\mathrm{corr}(X,Y) = \rho$, and $\underline{T}$ has a joint bivariate elliptical distribution with density function

$$C\ |V|^{-1/2}\ h(\underline{T}'V^{-1}\underline{T}) \qquad -\infty < X < \infty \qquad (2.1)$$
$$-\infty < Y < \infty,$$

where h is a differentiable function defined on $(-\infty,\infty)$, C is a normalizing constant, V is a 2×2 positive definite symmetric matrix.

A discrete observed random variate Z is defined by

$$Z = i \qquad \text{if } \alpha_i \le Y < \alpha_{i+1} \qquad (2.2)$$
$$i = 1,2,\ldots,r,$$

where α_i's are thresholds with $\alpha_1 = -\infty$ and $\alpha_{r+1} = \infty$ and r is the number of thresholds - 1.

Based on the observed random variables X and Z, we are going to develop an approach to estimate the unknown parameter vector $\underline{\theta}$ in the model which includes the thresholds and the polyserial correlation,

i.e. $\underline{\theta} = (\alpha_2,\alpha_3,\ldots\ldots,\alpha_r,\rho)'$.

Suppose we observe a random sample of $(X,Z)'$ of size N, where X is collected in continuous form, while Z is obtained in ordinal form and it is assumed that the variable vector $(X,Y)'$ is bivariate elliptically distributed with zero mean vector and covariance matrix V. The probability density function $P(x,z;\underline{\theta})$ of $(X,Z)'$ is a function of the unknown parameter vector $\underline{\theta}$. One of the main objectives of this thesis is to obtain an estimate of $\underline{\theta}$ by using the maximum likelihood method, and this ML estimate of $\underline{\theta}$ will be used to compare the estimate of $\underline{\theta}$ under the incorrect normal distribution assumption.

For the random vector $\underline{T} = (X,Y)'$, if it has an elliptical distribution with parameters $\underline{\mu} = (\mu_1, \mu_2)'$ and $V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}$, then corresponding characteristic function $\phi(t)$ is given as:

$$\phi(t) = E(\ \exp(it'\underline{T})\ )$$

$$= \exp(it'\underline{\mu})\ \Psi(t'Vt) \qquad \text{for some function } \Psi. \qquad (2.3)$$

The characteristic function of the marginal distribution of the random variable X can be obtained from $\phi(t)$ by putting $t = (t_1, 0)'$ and is shown as

$$\exp(it_1\mu_1)\ \Psi(t_1V_{11}t_1),$$

which is, in fact, the characteristic function of the random variable X with a univariate elliptical distribution. Therefore, marginal distribution of X is also elliptical with parameters μ_1 and

V_{11}. Since $\underline{T} = (X,Y)'$ has a bivariate elliptical distribution, it is well-known that the conditional distribution of Y given X is also elliptically distributed with parameters μ_x and ν_x given by

$$\mu_x = E(Y|X=x) = V_{21}\ V_{11}^{-1}\ x \qquad (2.4)$$

$$\text{and} \quad \nu_x = var(Y|X=x) = g(x)\ (V_{22} - V_{21}\ V_{11}^{-1}\ V_{21}) \qquad (2.5)$$

for some function g(.).

With these results, the probability density function of $(X,Z)'$, $P(x,z;\underline{\theta})$, can be decomposed into the marginal probability density function of X, $P_1(x;\mu_1,V_{11})$, and the probability of Z = z given that X = x, and the formulation looks like that

$$\begin{aligned} P(x,z;\underline{\theta}) &= P_1(x;\mu_1,V_{11}) \times Pr(Z=z|X=x) \\ &= P_1(x;\mu_1,V_{11}) \times Pr(\alpha_z \le Y < \alpha_{z+1}|X=x) \\ &= P_1(x;\mu_1,V_{11}) \times [Pr(Y < \alpha_{z+1}|X=x) - Pr(Y < \alpha_z|X=x)] \\ &= P_1(x;\mu_1,V_{11}) \times [\Omega(\alpha_{z+1}|X=x) - \Omega(\alpha_z|X=x)], \quad (2.6) \end{aligned}$$

where $\Omega(\alpha|X=x) = \int_{-\infty}^{\alpha} C\ \nu_x^{-1/2}\ h((Y-\mu_x)^2/\nu_x)\ dY.$

Thus, the likelihood function of the sample becomes

$$\begin{aligned} L(\underline{\theta}) &= P(x_1,z_1;\underline{\theta}) \times P(x_2,z_2;\underline{\theta}) \times \ldots\ldots.. \times P(x_N,z_N;\underline{\theta}) \\ &= \prod_{i=1}^{N} P_1(x_i;\mu_1,V_{11}) \times \prod_{i=1}^{N} [\Omega(\alpha_{z_i+1}|X=x_i) - \Omega(\alpha_{z_i}|X=x_i)]. \end{aligned}$$

Taking the natural logarithm, we have

$$\begin{aligned} Log\ L(\underline{\theta}) &= \sum_{i=1}^{N} Log\ P_1(x_i;\mu_1,V_{11}) \\ &+ \sum_{i=1}^{N} Log\ [\Omega(\alpha_{z_i+1}|X=x_i) - \Omega(\alpha_{z_i}|X=x_i)]. \end{aligned}$$

(2.7)

To obtain the maximum likelihood estimate of $\underline{\theta}$, $\underline{\theta}_{ML}$, we are required to solve the following system of equations.

$$\frac{\partial \operatorname{Log} L(\underline{\theta})}{\partial \alpha_j} = 0, \qquad j = 2,3,\ldots\ldots,r \tag{2.8}$$

and

$$\frac{\partial \operatorname{Log} L(\underline{\theta})}{\partial \rho} = 0. \tag{2.9}$$

In general, as Log $L(\underline{\theta})$ is a complicated nonlinear function of $\underline{\theta}$, the estimators cannot be solved in closed form, and so some iterative procedures have to be used to search for the estimates. In fact, since $L(\underline{\theta})$ is non-negative and the natural logarithm is a strictly increasing function, the maximization of $L(\underline{\theta})$ is equivalent to the minimization of $-\operatorname{Log} L(\underline{\theta})$, and a minimization subroutine written with FORTRAN IV is utilized to obtain the value of $\underline{\theta}$ such that $-\operatorname{Log} L(\underline{\theta})$ is minimized.

§2.2 Optimization Procedure

Of course, the generalized Newton-Raphson method (House-Holder,1953) has fast convergence, but it requires second derivatives of the objective function with respect to the parameter $\underline{\theta}$, and frequently fails to converge from a poor starting approximation to the minimum. Another powerful iterative method for finding a local minimum of a function of several variables is the Fletcher-Powell(1967) method.

The method is based on the Taylor series expansion. With the series expansion, the standard quadratic form of $f(\underline{\theta}) = -\text{Log } L(\underline{\theta})$ about the point $\underline{\theta}^i$ can be written as

$$f(\underline{\theta}) = f(\underline{\theta}^i) + (\underline{\theta} - \underline{\theta}^i)'g(\underline{\theta}^i) + 2^{-1}(\underline{\theta} - \underline{\theta}^i)'\, G^i\, (\underline{\theta} - \underline{\theta}^i)$$
$$+ E(\underline{\theta} - \underline{\theta}^i),$$

where $E(\underline{\theta} - \underline{\theta}^i)$ is the remainder, $\underline{g(\theta^i)}$ is the gradient vector at the point $\underline{\theta} = \underline{\theta}^i$, $\underline{\theta} = (\theta_1, \theta_2, \theta_3)'$, $G^i = (G^i_{kl})$, namely Hessian matrix and

$$G^i_{kl} = \left. \frac{\partial^2 f(\underline{\theta})}{\partial\theta_k\, \partial\theta_l} \right|_{\underline{\theta} = \underline{\theta}^i}.$$

When $\underline{\theta}$ is close to $\underline{\theta}^i$, the remainder can be ignored and we can expect that the quadratic form

$$Q(\underline{\theta}) = f(\underline{\theta}^i) + (\underline{\theta} - \underline{\theta}^i)'g(\underline{\theta}^i) + 2^{-1}(\underline{\theta} - \underline{\theta}^i)'\, G^i\, (\underline{\theta} - \underline{\theta}^i) \qquad (2.10)$$

will approximate $f(\underline{\theta})$. By differentiating equation (2.10) with respect to $\underline{\theta}$, the minimum point of $Q(\underline{\theta})$ can be given by the solution to the linear system of equations

$$g(\underline{\theta}^i) + G^i(\underline{\theta}_0 - \underline{\theta}^i) = 0,$$

$$\text{i.e. } \underline{\theta}_0 = \underline{\theta}^i - {G^i}^{-1} g(\underline{\theta}^i), \qquad (2.11)$$

where $\underline{\theta}_0$ is the estimate of parameter vector $\underline{\theta}$ at which $Q(\underline{\theta})$ is minimum. The equation (2.11) suggests the general iterative scheme

$$\underline{\theta}^{i+1} = \underline{\theta}^i - {G^i}^{-1} g(\underline{\theta}^i).$$

In the Fletcher-Powell method, the matrix ${G^i}^{-1}$ is not evaluated directly, instead, a positive definite symmetric matrix H is used which is initially chosen to be an identity matrix. This matrix will be modified after each iteration using the information gained by moving down the direction.

Let the current estimate be $\underline{\theta}^i$ with gradient $g(\underline{\theta}^i)$ and matrix H^i, the value of the estimate $\underline{\theta}^{i+1}$ at next step and the iteration algorithm can be stated as follows.

(1) Obtain the step size γ^i such that $f(\underline{\theta}^i + \gamma^i(-H^i g(\underline{\theta}^i)))$ is a minimum with respect to λ along $(\underline{\theta}^i + \lambda(-H^i g(\underline{\theta}^i)))$ by using a line search subroutine.

(2) Set $\underline{\theta}^{i+1} = \underline{\theta}^i + \gamma^i(-H^i g(\underline{\theta}^i))$.

(3) Evaluate $f(\underline{\theta}^{i+1})$ and $g(\underline{\theta}^{i+1})$ by objective function and gradient subroutine.

(4) Set $y(\underline{\theta}^{i}) = g(\underline{\theta}^{i+1}) - g(\underline{\theta}^{i})$.

(5) Update H by $H^{i+1} = H^{i} + A^{i} + B^{i}$, (2.12)

where $A^{i} = \dfrac{\sigma^{i}\,\sigma^{i\prime}}{\sigma^{i\prime}\,y(\underline{\theta}^{i})}$, $\sigma^{i} = \varphi^{i}(-H^{i}g(\underline{\theta}^{i}))$,

and $B^{i} = \dfrac{-H^{i}\,y(\underline{\theta}^{i})\,y(\underline{\theta}^{i})'\,H^{i}}{y(\underline{\theta}^{i})'\,H^{i}\,y(\underline{\theta}^{i})}$.

(6) Set i = i + 1 and repeat above procedure.

The procedure is terminated when every component of $-H^{i}g(\underline{\theta}^{i})$ is less than a prescribed accuracy level.

In the above procedure, it is neccessary to find the first partial derivative of Log $L(\underline{\theta})$ with respect to $\underline{\theta}$(i.e. namely, the gradient vector), which is, in turn, a function of the first partial derivative of $\Omega(\alpha|X=x)$ with respect to $\underline{\theta}$. For different members of elliptical distribution family, the exact expressions of $\partial\Omega(\alpha|X=x)/\partial\underline{\theta}$ are also different, but under some mild conditions on differentiability and integrability, by the theorem of basic calculus, it can be shown that

if $\Omega(\alpha|X=x) = \int_{-\infty}^{\alpha} C\,\nu_{x}^{-1/2}\,h((Y-\mu_{x})^{2}/\nu_{x})\,dY$,

where $\mu_{x} = V_{21}\,V_{11}^{-1}\,x$,

and $\nu_{x} = g(x)\,(V_{22} - V_{21}\,V_{11}^{-1}\,V_{21})$, then $$\frac{\partial\, \Omega(\alpha|X=x)}{\partial\, \alpha} = C\, \nu_x^{-1/2}\, h((\alpha-\mu_x)^2/\nu_x).$$

This expression is only a univariate function in X. Also,

$$\frac{\partial\, \Omega(\alpha|X=x)}{\partial\, \rho} = \frac{\partial \int_{-\infty}^{\alpha} C\, \nu_x^{-1/2}\, h((Y-\mu_x)^2/\nu_x)\, dY}{\partial\, \rho}$$

$$= \frac{\partial \int_{-\infty}^{\alpha^*} C\, h(\omega^2)\, d\omega}{\partial\, \rho}$$

$$= C\, h(\alpha^{*2})\, \frac{\partial \alpha^*}{\partial \rho}, \qquad (2.13)$$

where $\omega = (Y-\mu_x)\, \nu_x^{-1/2}$, $\quad \alpha^* = (\alpha-\mu_x)\, \nu_x^{-1/2}$.

The above expressions give the general formula for computing the gradient vector of the objective function in the context of the bivariate elliptical distribution family. Although, the expression of $\partial\alpha^*/\partial\rho$ are different for different members of elliptical distribution family, it can be easily deduced exactly for each one.

In the subsequent sections, we will pay more attention on two members of the bivariate elliptical distribution family, namely, the elliptical t distribution and the contaminated normal distribution.

§2.3 Elliptical t distribution

For a m-dimension random vector $\underline{T}$, if it is elliptically t distributed with n degree of freedom, then $\underline{T}$ has the following probability density function

$$\frac{\Gamma[(n+m)/2]}{\Gamma(n/2)\ (n\pi)^{m/2}}\ |V|^{-1/2}\ [1 + n^{-1}\ (\underline{T} - \underline{\mu})'V^{-1}(\underline{T} - \underline{\mu})]^{-(n+m)/2}, \qquad (2.14)$$

where $\Gamma(.)$ is the usual Gamma function.

For a bivariate elliptical t distribution (i.e. m=2), the random vector $\underline{T} = (X,Y)'$ with $\underline{\mu} = \underline{0}$, and $V = \frac{n-2}{n}\begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$, $n > 2$, is selected so as to simplify subsequent derivation. The corresponding probability density function become

$$\begin{aligned} t_2(X,Y;V,n) &= \frac{\Gamma(1+n/2)}{\Gamma(n/2)\ (n\pi)}\ |V|^{-1/2}\ [1+n^{-1} \times \underline{T}'V^{-1}\underline{T}]^{-(n+2)/2} \\ &= \frac{1}{2\pi} \times |V|^{-1/2}[1+n^{-1} \times \underline{T}'V^{-1}\underline{T}]^{-(n+2)/2}, \qquad (2.15) \end{aligned}$$

where $\underline{T} = (X,Y)'$ with $-\infty < X < \infty$, $-\infty < Y < \infty$.

To obtain the exact expression of the partial derivatives $\partial \text{Log } L(\underline{\theta})/\partial \alpha_j$, $j=2,\dots,r$, and $\partial \text{Log } L(\underline{\theta})/\partial \rho$, we denote the probability density function of the univariate t distribution with parameters μ and ν by $t_1(Y;\mu,\nu,n)$, and its expression is given as belows.

$$t_1(Y;\mu,\nu,n) = \frac{\Gamma[(n+1)/2]}{\Gamma(n/2)\sqrt{n\pi}} \frac{1}{\sqrt{\nu}} [\,1 + (Y-u)^2/(n\nu)]^{-(n+1)/2}, \tag{2.16}$$

where n is the degree of freedom and $-\infty < Y < \infty$.

It can be shown that (details of the proof are given in Appendix A)

$$t_2(X,Y;V,n) = t_1(X;0,(n-2)/n,n)$$
$$\times\ t_1(Y;\rho x, \frac{n-2}{n+1} \times (1 + \frac{x^2}{n-2}) \times (1-\rho^2), n+1), \tag{2.17}$$

where $t_1(x;0,(n-2)/n,n)$ is the marginal density function of X and $t_1(Y;\rho x, \frac{n-2}{n+1} \times (1 + \frac{x^2}{n-2}) \times (1-\rho^2), n+1)$ is the conditional density function of Y given X=x.

Similar to the previous derivation in the general elliptical distribution family, the joint probability density function of $(X,Z)'$ is

$$P(x,z;\underline{\theta}) = t_1(x;0,(n-2)/n,n) \times [T(\alpha_{z+1}|X=x) - T(\alpha_z|X=x)], \tag{2.18}$$

where $T(\alpha_z|X=x) = \int_{-\infty}^{\alpha_z} t_1(Y;\rho x,\frac{n-2}{n+1}(1 + \frac{x^2}{n-2})(1-\rho^2),n+1)\, dY$.

In fact, $T(\alpha_z|X=x)$ is the univariate t distribution function. For evaluating the probability integrals of $T(\alpha_z|X=x)$, we can make use of the following equality

$$T(\alpha_z|X=x) = \int_{-\infty}^{\alpha_z} t_1(Y;\mu_x,\nu_x,n+1)\, dY$$

$$= \int_{-\infty}^{\alpha_z^*} t_1(\omega;0,1,n+1)\, d\omega, \qquad (2.19)$$

where $\mu_x = \rho x$, $\nu_x = \frac{n-2}{n+1}\left(1 + \frac{x^2}{n-2}\right)(1-\rho^2)$,

$\omega = (Y-\mu_x)\nu_x^{-1/2}$ and $\alpha_z^* = (\alpha_z-\mu_x)\nu_x^{-1/2}$.

The above equality can be easily proved through the change of variable, and $T(\alpha_z|X=x)$ can be evaluated with high accuracy by using the computing algorithm published by Cooper(1968). Furthermore, the natural logarithm of the maximum likelihood function become

$$\text{Log } L(\underline{\theta}) = \sum_{i=1}^{N} \text{Log } t_1(x_i;0,\frac{n-2}{n},n) + \sum_{i=1}^{N} \text{Log}[T(\alpha_{z_i+1}|X=x_i)-T(\alpha_{z_i}|X=x_i)]. \qquad (2.20)$$

Since the first term on the right hand side of above expression (2.20) is independent of α_j and ρ, the partial derivative of this term with respect to them will be zero, and the expression of partial derivatives $\frac{\partial \text{ Log } L(\underline{\theta})}{\partial\alpha_j}$, $j=2,\ldots,r$ and $\frac{\partial \text{ Log } L(\underline{\theta})}{\partial \rho}$ can be found as belows.

$$\frac{\partial \text{ Log } L(\underline{\theta})}{\partial \alpha_j} = \frac{\partial \sum_{i=1}^{N} \text{Log } [T(\alpha_{z_i+1}|X=x_i)-T(\alpha_{z_i}|X=x_i)]}{\partial \alpha_j}$$

$$= \sum_{i=1}^{N} \frac{1}{T(\alpha_{z_i+1}|X=x_i)-T(\alpha_{z_i}|X=x_i)} \frac{\partial[T(\alpha_{z_i+1}|X=x_i)-T(\alpha_{z_i}|X=x_i)]}{\partial \alpha_j}. \qquad (2.21)$$

Since α_k is independent of α_j for $k\neq j$, the partial derivative of $\dfrac{\partial\, T(\alpha_k|X=x_i)}{\partial\, \alpha_j}$ will be zero for k≠j. In addition, when k=j, the partial derivative $\dfrac{\partial\, T(\alpha_k|X=x_i)}{\partial\, \alpha_j}$ become

$$\frac{\partial\, T(\alpha_j|X=x_i)}{\partial\, \alpha_j} = \frac{\int_{-\infty}^{\alpha_j} t_1(Y;\mu_x,\nu_x,n+1)\, dY}{\partial\, \alpha_j}$$
$$= t_1(\alpha_j;\mu_x,\nu_x,n+1),$$

where

$$t_1(\alpha_j;\mu_x,\nu_x,n+1) = \frac{\Gamma[(n+2)/2]}{\Gamma[(n+1)/2]\cdot\sqrt{(n+1)\pi}}\,\frac{1}{\sqrt{\nu_x}}\,[1+\frac{(\alpha_j-\mu_x)^2}{(n+1)\nu_x}]^{-\frac{n+2}{2}}. \tag{2.22}$$

As a result, the complete expression of (2.21) is given as

$$\frac{\partial\, \text{Log } L(\underline{\theta})}{\partial\, \alpha_j} = \sum_{i=1}^{N} \frac{1}{T(\alpha_{z_i+1}|X=x_i)-T(\alpha_{z_i}|X=x_i)}$$
$$\times [\frac{\partial\, T(\alpha_{z_i+1}|X=x_i)}{\partial\alpha_j} - \frac{\partial\, T(\alpha_{z_i}|X=x_i)}{\partial\alpha_j}],$$

where $\dfrac{\partial\, T(\alpha_{z_i+1}|X=x_i)}{\partial\alpha_j} = \begin{cases} t_1(\alpha_j;\mu_x,\nu_x,n+1) & \text{if } z_i = j-1 \\ 0 & \text{otherwise,} \end{cases}$

and $\dfrac{\partial\, T(\alpha_{z_i}|X=x_i)}{\partial\alpha_j} = \begin{cases} t_1(\alpha_j;\mu_x,\nu_x,n+1) & \text{if } z_i = j \\ 0 & \text{otherwise.} \end{cases}$

Similarly, with suitable transformation, the expression for the first partial derivative of Log $L(\underline{\theta})$ with respect to ρ can be derived as follows(refer Appendix B for the details of the proof).

$$\frac{\partial\, \text{Log } L(\underline{\theta})}{\partial\, \rho} = \frac{\partial \sum_{i=1}^{N} \text{Log}[T(\alpha_{z_i+1}|X=x_i) - T(\alpha_{z_i}|X=x_i)]}{\partial\, \rho}$$

$$= \sum_{i=1}^{N} \frac{1}{T(\alpha_{z_i+1} \mid X=x_i) - T(\alpha_{z_i} \mid X=x_i)} \times$$

$$\left\{ \frac{\partial \int_{-\infty}^{\alpha_{z_i+1}} t_1(Y;\mu_X,\nu_X,n+1)\, dY}{\partial \rho} - \frac{\partial \int_{-\infty}^{\alpha_{z_i}} t_1(Y;\mu_X,\nu_X,n+1)\, dY}{\partial \rho} \right\}$$

$$= \sum_{i=1}^{N} \frac{1}{T(\alpha_{z_i+1} \mid X=x_i) - T(\alpha_{z_i} \mid X=x_i)} \times$$

$$\left\{ \frac{t_1(\alpha^*_{z_i+1};0,1,n+1)(\rho\cdot\alpha_{z_i+1} - x_i) - t_1(\alpha^*_{z_i};0,1,n+1)(\rho\cdot\alpha_{z_i} - x_i)}{\sqrt{\nu_X}\cdot(1-\rho^2)} \right\}. \tag{2.23}$$

§2.4 Contaminated Normal Distribution

Now, We will look into the details in the case of the bivariate elliptical contaminated normal distribution, the probability density function with parameters ϵ and σ is given as

$$g_2(\underline{T};\epsilon,\sigma) = \frac{(1-\epsilon)\ |V|^{-1/2}\ \exp(-2^{-1}\ \underline{T}'V^{-1}\underline{T})}{2\pi}$$

$$+ \frac{\epsilon\ |\sigma^2 V|^{-1/2}\ \exp(-2^{-1}\sigma^{-2}\ \underline{T}'V^{-1}\underline{T})}{2\pi}, \qquad (2.24)$$

where $\underline{T} = (X,Y)'$, $V = \frac{1}{C}\begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}$, $C = (1-\epsilon) + \sigma^2\epsilon$ and $0 < \epsilon < 1$, $\sigma > 0$. In fact, it is the sum of the density function of two normal distributions, namely, $N_2(\underline{0},V)$ and $N_2(\underline{0},\sigma^2 V)$. In order to obtain the partial derivative of the corresponding Log-likelihood function with respect to α_j and ρ, we can make use of the results of conditional distribution of Y given X=x, and decompose the density functions of the distribution $N_2(\underline{0},V)$ and $N_2(\underline{0},\sigma^2 V)$ into products of their marginal density functions and their conditional density functions. Thus, $g_2(\underline{T};\epsilon,\sigma)$ can be written as

$$g_2(\underline{T};\epsilon,\sigma) = \frac{(1-\epsilon)\sqrt{C}}{\sqrt{2\pi}}\exp(-CX^2\cdot 2^{-1})\ \frac{1}{\sqrt{2\pi}\ \sqrt{\nu_x}}\ \exp(-(Y-\mu_x)^2\cdot(2\nu_x)^{-1})$$

$$+ \frac{\epsilon\ \sqrt{C}}{\sqrt{2\pi}\ \sigma}\exp(-CX^2\cdot 2^{-1}\sigma^{-2})\ \frac{1}{\sqrt{2\pi}\ \sigma\ \sqrt{\nu_x}}\ \exp(-(Y-\mu_x)^2\cdot(2\sigma^2\nu_x)^{-1}), \qquad (2.25)$$

where $\mu_x = V_{21}\ V_{11}^{-1}\ x = \rho\ x,$ (2.26)

and $\nu_x = V_{22} - V_{21}\ V_{11}^{-1}\ V_{12} = (1-\rho^2)/C.$ (2.27)

Let $g_1(Y;\mu,\nu)$ be the normal density function with mean μ and variance ν,

i.e. $g_1(Y;\mu,\nu) = \dfrac{1}{\sqrt{2\pi\nu}} \exp(-(Y-\mu)^2\cdot(2\nu)^{-1})$.

Then, the joint probability density function of $(X,Z)'$ can be written as

$$P(x,z;\underline{\theta}) = (1-\epsilon)g_1(x;0,1/c)\times[\phi_1(\alpha_{z+1};\mu_x,\nu_x|X=x)-\phi_1(\alpha_z;\mu_x,\nu_x|X=x)]$$
$$+ \epsilon g_1(x;0,\sigma^2/c)\times[\phi_1(\alpha_{z+1};\mu_x,\sigma^2\nu_x|X=x)-\phi_1(\alpha_z;\mu_x,\sigma^2\nu_x|X=x)], \quad (2.28)$$

where $\phi_1(\alpha_z;\mu_x,\nu_x|X=x)$ is the univariate normal distribution function with mean μ_x and variance ν_x,

i.e. $\phi_1(\alpha_z;\mu_x,\nu_x|X=x) = \displaystyle\int_{-\infty}^{\alpha_z} g_1(Y;\mu_x,\nu_x)\, dY$.

Similar to the case of elliptical t distribution(see (2.19)), taking the transformation

$\omega = (Y-\mu_x)\nu_x^{-1/2}$ and $\alpha_z^* = (\alpha_z-\mu_x)\nu_x^{-1/2}$,

we have

$$\phi_1(\alpha_z;\mu_x,\nu_x|X=x) = \phi_1(\alpha_z^*;0,1|X=x) \quad (2.29)$$

and so $P(x,z;\underline{\theta}) = (1-\epsilon)g_1(x;0,1/c)\times[\phi_1(\alpha_{z+1}^*;0,1|X=x)-\phi_1(\alpha_z^*;0,1|X=x)]$
$+ \epsilon\, g_1(x,0,\sigma^2/c)\times[\phi_1(\alpha_{z+1}^+;0,1|X=x)-\phi_1(\alpha_z^+;0,1|X=x)]$,

where $\alpha_z^+ = \dfrac{\alpha_z - \mu_x}{\sigma\sqrt{\nu_x}}$.

The natural logarithm of the maximum likelihood function Log

$L(\underline{\theta})$ become

$$\sum_{i=1}^{N} \text{Log}\{(1-\epsilon)g_1(x_i;0,1/c)\times[\phi_1(\alpha^*_{z_i+1};0,1|X=x_i)-\phi_1(\alpha^*_{z_i};0,1|X=x_i)]$$
$$+\ \epsilon g_1(x_i;0,\sigma^2/c)\times[\phi_1(\alpha^+_{z_i+1};0,1|X=x_i)-\phi_1(\alpha^+_{z_i};0,1|X=x_i)]\}. \qquad (2.30)$$

On differentiating the above expression (2.30) with respect to α_j, $j=2,3,\ldots,r$, we obtain the expression

$$\frac{\partial \text{ Log } L(\underline{\theta})}{\partial \alpha_j} = \sum_{i=1}^{N} \frac{1}{D}\Big\{(1-\epsilon)g_1(x_i;0,\frac{1}{c})\Big[\frac{\partial\phi_1(\alpha^*_{z_i+1};0,1|X=x_i)}{\partial \alpha_j}$$
$$-\frac{\partial\phi_1(\alpha^*_{z_i};0,1|X=x_i)}{\partial\alpha_j}\Big] + \epsilon\ g_1(x_i;0,\frac{\sigma^2}{C})\Big[\frac{\partial\phi_1(\alpha^+_{z_i+1};0,1|X=x_i)}{\partial \alpha_j}$$
$$-\frac{\partial\phi_1(\alpha^+_{z_i};0,1|X=x_i)}{\partial\alpha_j}\Big]\Big\}, \qquad (2.31)$$

where $D = (1-\epsilon)\ g_1(x_i;0,\frac{1}{C}) \times [\phi_1(\alpha^*_{z_i+1};0,1|X=x_i) - \phi_1(\alpha^*_{z_i};0,1|X=x_i)]$

$$+\quad \epsilon\ g_1(x_i;0,\frac{\sigma^2}{C}) \times [\phi_1(\alpha^+_{z_i+1};0,1|X=x_i) - \phi_1(\alpha^+_{z_i};0,1|X=x_i)],$$

$$\frac{\partial\phi_1(\alpha^*_{z_i+1};0,1|X=x_i)}{\partial \alpha_j} = \begin{cases} g_1(\alpha_j;\mu_x,\nu_x) & \text{if } z_i=j-1 \\ 0 & \text{otherwise,} \end{cases}$$

$$\frac{\partial\phi_1(\alpha^*_{z_i};0,1|X=x_i)}{\partial\alpha_j} = \begin{cases} g_1(\alpha_j;\mu_x,\nu_x) & \text{if } z_i=j \\ 0 & \text{otherwise,} \end{cases}$$

$$\frac{\partial\phi_1(\alpha^+_{z_i+1};0,1|X=x_i)}{\partial \alpha_j} = \begin{cases} g_1(\alpha_j;\mu_x,\sigma^2\nu_x) & \text{if } z_i=j-1 \\ 0 & \text{otherwise,} \end{cases}$$

and $$\frac{\partial\phi_1(\alpha^+_{z_i};0,1|X=x_i)}{\partial\alpha_j} = \begin{cases} g_1(\alpha_j;\mu_x,\sigma^2\nu_x) & \text{if } z_i=j \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, the expression of $\dfrac{\partial \text{ Log } L(\underline{\theta})}{\partial \rho}$ can be obtained as follows.

$$\frac{\partial \text{ Log } L(\underline{\theta})}{\partial \rho}$$

$$= \sum_{i=1}^{N} \frac{1}{D}\left\{ (1-\epsilon)\, g_1(x_i;0,\tfrac{1}{c}) \left[\frac{\partial\phi_1(\alpha^*_{z_i+1};0,1|X=x_i)}{\partial \rho} - \frac{\partial\phi_1(\alpha^*_{z_i};0,1|X=x_i)}{\partial \rho}\right]\right.$$

$$\left. + \ \epsilon \ \ g_1(x_i;0,\tfrac{\sigma^2}{c}) \left[\frac{\partial\phi_1(\alpha^+_{z_i+1};0,1|X=x_i)}{\partial \rho} - \frac{\partial\phi_1(\alpha^+_{z_i};0,1|X=x_i)}{\partial \rho}\right]\right\}. \tag{2.32}$$

where
$$\frac{\partial\phi_1(\alpha^*_{z_i+1};0,1|X=x_i)}{\partial \rho} = \frac{\partial\phi_1(\alpha^*_{z_i+1};0,1|X=x_i)}{\partial\alpha^*_{z_i+1}}\ \frac{\partial\alpha^*_{z_i+1}}{\partial \rho}$$

$$= \frac{\sqrt{c}}{\sqrt{2\pi}} \exp(-\tfrac{1}{2}\alpha^{*\,2}_{z_i+1}) \times \frac{\rho\,\alpha_{z_i+1} - x_i}{(1-\rho^2)^{3/2}} .$$

$$\frac{\partial\phi_1(\alpha^*_{z_i};0,1|X=x_i)}{\partial \rho} = \frac{\sqrt{c}}{\sqrt{2\pi}} \exp(-\tfrac{1}{2}\alpha^{*\,2}_{z_i}) \times \frac{\rho\,\alpha_{z_i} - x_i}{(1-\rho^2)^{3/2}},$$

$$\frac{\partial\phi_1(\alpha^+_{z_i+1};0,1|X=x_i)}{\partial \rho} = \frac{\sqrt{c}}{\sqrt{2\pi}} \exp(-\tfrac{1}{2}\alpha^{+\,2}_{z_i+1}) \times \frac{\rho\,\alpha_{z_i+1} - x_i}{\sigma\ (1-\rho^2)^{3/2}}.$$

and
$$\frac{\partial\phi_1(\alpha^+_{z_i};0,1|X=x_i)}{\partial \rho} = \frac{\sqrt{c}}{\sqrt{2\pi}} \exp(-\tfrac{1}{2}\alpha^{+\,2}_{z_i}) \times \frac{\rho\,\alpha_{z_i} - x_i}{\sigma\ (1-\rho^2)^{3/2}} .$$

Chapter 3 Technical Details

§3.1 Structure of the program

By using the results developed in Chapter 2, a computer program written in FORTRAN IV is implemented to estimate the parameter vector $(\alpha_2, \alpha_3, \ldots \alpha_r, \rho)'$.

We are going to study the estimation of polyserial correlation and thresholds of two special members of elliptical distribution family, namely, elliptical t and contaminated normal distribution. The robustness of estimators proposed by Poon and Lee(1987) against the normal assumption will be studied also. As a result, four sets of programs are prepared, and their structure are similar. According to functions, the subroutines in the programs can be classified into four groups. The first group composes of subroutines which are used to generate sample variates of elliptical distribution family. They include uniform random number, bivariate standard normal distribution and elliptical distribution generators.

The second group of subroutines are developed especially for the two elliptical distributions. These subroutines are used to calculate values of the objective function, and compute its first partial derivative with respect to parameters. They also involve some function subroutines used to compute cumulative distribution function values and density function values of univariate t-distribution and univariate normal distribution. The third group is the heart of the program, its job mainly deals with minimization of the objective function, so it includes line search subroutine and Fletcher-Powell algorithm. The fourth group consists of utility programs which serve various purposes including computation of sample correlation coefficients, means and root mean squares of the estimates. A main program coordinates the operation of every subroutine, controls the input and output information, sets the tolerance level of convergence and sets the parameters used to generate sample variates. In addition, the subroutines from each group are linked together by the main program to form one set of programs.

§3.2 Input and output of the program

To execute the programs, some information such as degree of freedom, ϵ, σ and sample size must be supplied by the user. These data will be used by the program as references to generate the sample variates and to estimate the parameters. For elliptical t distribution, the following data must be supplied before executing the programs.

(1) Sample Size N.

(2) Degree of freedom.

(3) The tolerance level of convergence.

(4) The number of replication.

(5) Maximum number of iteration allowed.

For the contaminated normal distribution, similar data are supplied except the degree of freedom. In addition, the parameter ϵ and σ of contaminated normal distribution should be supplied in advance.

To maintain high degree of accuracy of the estimation, in handling the problem of infinity, a large number will not be used to provide a substitute for infinity. In fact, for univariate t distribution and standard normal distribution, the values of cumulative distribution functions at negative infinity, $T(-\infty|X=x)$ and $\phi_1(-\infty;\mu_X,\nu_X|X=x)$, are substituted by zero, the values of these functions at positive infinity, $T(\infty|X=x)$ and $\phi_1(\infty;\mu_x,\nu_x|X=x)$, are substituted by one, and their corresponding density function at both infinities are substituted by zero. These substitions will not only avoid overflow and underflow problems, but also maintain the computation accuracy.

Apart from echoes of the user's supplied information, the computer output also contains values of the estimates of the thresholds and polyserial correlation in each replication. The mean of 50 estimates for each parameter and its root mean square is printed as well.

For detail investigation of each replication, the program can provide another form of output format. In detail output format, it gives information such as the values of initial estimates, user's supplied information and the number of iteration for convergence. Besides, in each iteration, the values of new estimates, the gradient vector and the H matrix described in (2.12) are printed for references. As the aim of the optimization procedure is to minimize the objective function, the value of the objective function in each iteration is also given for checking convergency of the procedure.

§3.3 Methods for obtaining initial estimates

As the minimum of the objective function cannot be solved algebraically in closed form, nonlinear optimization procedure is required to compute the optimum estimate. From the nature of the objective function, we can see that it is very difficult to derive its Hessian matrix. Hence, it is not worth while to apply the classical Newton-Raphson algorithm.

In this thesis, the Fletcher-Powell algorithm is used because it only requires the gradient vector of the objective function, an initial estimate of $\underline{\theta}$ and an initial positive definite matrix. This initial positive definite matrix can be arbitrarily chosen, thus the identity matrix will be used for simplicity. This matrix will be updated in each iteration to approximate the inverse of the Hessian matrix.

In general, the Fletcher-Powell algorithm is very robust to the starting values of the estimates. However, experience indicates that a good starting value would reduce the time of convergence. Therefore, a sample estimate that is based on $(x,z)'$, $i=1,2,\ldots,N$, is used to estimate the initial value of $\underline{\theta}$. In our studies, the initial estimate of ρ used by Poon & Lee(1987) is utilized. That is, the sample product correlation coefficient of X and Z is taken as an initial estimate of ρ.

The starting values for α_i, $i=2,3,\ldots,r$, are obtained from the cumulative marginal proportions. That is,

$$\hat{\alpha}_k = \xi^{-1}(p_k), \quad k=2,3,\ldots,r \quad \text{with } \hat{\alpha}_1 = -\infty \text{ and } \hat{\alpha}_{r+1} = \infty. \tag{3.1}$$

where ξ^{-1} denotes the inverse marginal distribution function of elliptical t or contaminated normal distribution, and p_k is given as

$$p_k = \frac{\text{total number of samples whose z values are less than k}}{\text{total number of samples}} \tag{3.2}$$

For user's convenience, there is no need for user to input the initial values of the estimates. Instead, they are automatically computed by the program. As the initial estimate of ρ, the sample product correlation coefficient can be readily obtained by a subroutine based on the values of $(x,z)'$.

For obtaining initial estimates of the thresholds $\hat{\alpha}_k$, more work have to be done. Because the marginal distribution function ξ of a bivariate elliptical t distribution is not a standard t distribution, transformation is needed before utilizing a subroutine to compute the value of the inverse marginal t distribution function. Since the initial estimate $\hat{\alpha}_k$, $k = 2,3,\ldots,r$, can be considered as the $100p_k$ percentiles of the marginal cumulative distribution function of Y. That is,

$$p_k = \int_{-\infty}^{\hat{\alpha}_k} t_1(Y;0,\frac{n-2}{n},n)dY, \qquad (3.3)$$

where $t_1(Y;0,\frac{n-2}{n},n)$ is the marginal density function of Y (see (2.17)). By making transformation $\omega = Y\cdot(\frac{n-2}{n})^{-1/2}$, it can be shown that

$$p_k = \int_{-\infty}^{\varsigma(\hat{\alpha}_k)} t_1(\omega;0,1,n)d\omega, \qquad (3.4)$$

where $\varsigma(\hat{\alpha}_k) = \hat{\alpha}_k\cdot(\frac{n-2}{n})^{-1/2}$.

Therefore, by using the expression (3.2), the value of P_k can be easily obtained from the samples. In addition, with the algorithm developed by Hill(1970), $\varsigma(\hat{\alpha}_k)$ can be computed from the expression (3.4). Finally, just multiplying $\varsigma(\hat{\alpha}_k)$ obtained in the algorithm by $(\frac{n-2}{n})^{1/2}$, then initial threshold estimate $\hat{\alpha}_k$ is obtained.

Furthermore, under the assumption of the contaminated normal distribution, we can realize that as the marginal distribution of Y is a mixture of two univariate normal distribution, namely, $N_1(0,\frac{1}{c})$ and $N_1(0,\frac{\sigma^2}{c})$, combined in the ratio $(1-\epsilon)$ and ϵ, where c is defined in the expression (2.24) as $c = (1-\epsilon) + \sigma^2\epsilon$, the inverse marginal distribution function is difficult to find. However, in general, ϵ is usually small and σ is also not too large. In order to speed up the optimization procedure and reduce the time of estimation in each replication, the percentile value of $N_1(0,\frac{1}{c})$ will be used to approximate the correct percentile of the contaminated normal distribution.

From the expression (2.25), the marginal density function of Y can be given as

$$g_1(Y;0,\frac{1}{c}) = \frac{\sqrt{c}}{\sqrt{2\pi}} \exp(-\frac{c}{2}Y^2). \qquad (3.5)$$

Similarly, the initial estimate $\hat{\alpha}_k$ can be regarded as the $100p_k$ percentile of the marginal cumulative distribution function of Y,

i.e. $p_k = \int_{-\infty}^{\hat{\alpha}_k} g_1(Y;0,\frac{1}{c})dY.$

With the same transformation as before, we have

$$p_k = \int_{-\infty}^{\hat{\alpha}_k\sqrt{c}} g_1(\omega;0,1)\, d\omega$$

$$= \Phi_1(\hat{\alpha}_k\sqrt{c}), \qquad (3.6)$$

where $\omega = Y\sqrt{c}$ and $\Phi_1(\hat{\alpha}_k\sqrt{c})$ is the cumulative distribution function of the standard univariate normal distribution. Thus, the starting values can be obtained by the following expression:

$$\hat{\alpha}_k = \frac{1}{\sqrt{c}} \Phi_1^{-1}(p_k), \qquad k = 2,3,\ldots,r. \qquad (3.7)$$

As a result, the algorithm written by Beasley and Springer(1977) which computes the $100p_k$ percentile of standard normal distribution can be utilized to serve this purpose.

Chapter 4 Simulation Study

§4.1 Outline of the study

In this thesis, two sets of simulation studies are carried out separately to investigate the performance and the behaviour of the estimates from the maximum likelihood procedure under two elliptical distributions with various combination of parameters. In addition, two sets of studies will be conducted to exploit the effect on the performance and the behaviour of the estimate when the estimation is under incorrect normality assumption, while the sample variates are still come from these two bivariate elliptical distributions. In order to make comparison, the estimation procedures are executed repeatly under the same environment. That is, all the default values used in the program are the same except the parameters and the distribution assumption in which we are interested. Moreover, the Monto Carlo data are also generated by the same set of seed numbers.

Except the special parameters concerned with separate particular bivariate elliptical distribution, e.g. degree of freedom, ϵ and σ, other different combination of parameter values are the same, e.g. thresholds and correlation ρ. Furthermore, these values will be chosen such that the samples may have different configurations of distribution. In our studies, the polyserial correlations ρ are chosen at ρ = 0.0, 0.2, 0.5 and 0.7. This means that the samples of X and Y varies from weakly correlated to strongly correlated. In addition, the thresholds are also selected such that the marginal distribution of Z change from positively skewed to negatively skewed, and accordingly, the values of thresholds are selected as the following.

(1) $\alpha_1 = -\infty$, $\alpha_2 = -1.0$, $\alpha_3 = 0.0$, $\alpha_4 = \infty$.

(2) $\alpha_1 = -\infty$, $\alpha_2 = -0.5$, $\alpha_3 = 0.5$, $\alpha_4 = \infty$.

(3) $\alpha_1 = -\infty$, $\alpha_2 = -0.3$, $\alpha_3 = 0.7$, $\alpha_4 = \infty$.

From the above different combinations of thresholds, we can clearly see that r is set to 3, values of α_1 and α_4 are negative infinity and positive infinity respectively, and so we are only interested in estimating and exploiting the behaviour of the thresholds α_2 and α_3. Similarly, three values of sample size are chosen, namely, small(N=20), moderate(N=40) and large(N=100).

For other special parameters, they are designed for each member of elliptical distribution family. As an important parameter of t distribution, degree of freedom must be given in order to describe the complete distribution, and the values of degree of freedom are assigned such that the deviation from normality distribution ranges from large to moderate. Thus, in our studies, they take the values n=5 and 10.

Similarly, for the contaminated normal distribution, we need both parameters ϵ and σ to generate the desired distributions. ϵ = 0.1, 0.3 and 0.5, and σ = 0.1 and 0.5 are chosen in our studies.

For the elliptical t distribution, since two different values of degree of freedom(i.e. d.f. = 5 and 10), three different values of sample size(i.e. N = 20, 40 and 100), three different types of thresholds configuration and four different values of polyserial correlation(i.e. ρ = 0.0, 0.2, 0.5 and 0.7) are selected in the simulation studies, total 72(i.e. 2 x 3 x 3 x 4) combinations will be considered, and each combination includes 50 replications of estimation for the parameters α_2, α_3 and ρ. Similarly, for the contaminated normal distribution, same set of sample size, thresholds and polyserial correlation are used. However, instead of the degree of freedom, three different values of ϵ(i.e. ϵ = 0.1, 0.3 and 0.5) and two different values of σ(i.e. σ = 0.1 and 0.5) are used in the simulation studies. As a result, there are 216(i.e. 3 x 3 x 4 x 3 x 2) combinations to be investigated, and also each one involves 50 replications of the estimation for the parameters α_2, α_3 and ρ.

For every 50 replications, the means of the estimates and the root mean squares are computed as follows.

$$\text{mean} = \frac{1}{50} \sum_{k=1}^{50} \hat{\theta}_k \, , \tag{4.1}$$

$$RMS = \left[\frac{1}{50} \sum_{k=1}^{50} (\hat{\theta}_k - \theta)^2 \right]^{1/2}, \qquad (4.2)$$

where $\hat{\theta}_k$ denotes the estimate of α_2, α_3 and ρ obtained from the k^{th} replication and θ denotes the true value of the parameter α_2, α_3 and ρ.

The RMS so computed is used for assessing the bias of the estimates, and the mean is a representative of all estimates obtained in 50 replications. Under one particular set of elliptical distribution and parameters α_2, α_3 and ρ, the values of RMS and means are presented with their corresponding counterparts in table 4.1 to table 4.12 (see table section of this thesis). Apart from using the different combination of parameters such as degree of freedom, ϵ and σ in estimating the parameters α_2, α_3 and ρ, the estimation procedure in the program also include the following default values as references in estimation.

(1) The maximum number of iterations allowed,
(50 is set in our studies).

(2) The tolerance level of convergence for $-\ddot{H}^i g(\theta^i)$ (0.001 is set).

These values can be changed easily by user if neccessary. In addition, every component of $-\ddot{H}\cdot g(\underline{\theta})$, which has been described in Chapter 2, is a measure to determine whether the optimization procedure converges or not. That is, if every component is less than the prescribed tolerance level for convergence, then the procedure is declared to be convergent, and the most updated estimates will be considered to be the estimates of the parameters in that replication. Otherwise, if the convergence criterion is not satisfied beyond the maximum number of iterations allowed, then the procedure is sentenced to be divergent, execution of the optimization procedure will terminate at once, and the next new replication begin until completing all 50 replications.

4.2 Generations of data

§4.2.1 elliptical t distribution

To generate the bivariate elliptical t variates, we can make use of the following results given in Muirhead(1982, p.48) and the derivation is given in Appendix C.

For $n > 2$,

$$\underline{T} = \sqrt{n}\, U^{-1/2}\, V^{1/2}\, \underline{X} \qquad (4.3)$$

has a bivariate elliptical t distribution, where $\underline{X}$ is distributed as $N_2(\underline{0}, I_2)$, U is distributed as χ^2_n and are stochastically independent of $\underline{X}$, and $V^{1/2}$ is a symmetric square root of a 2×2 positive definite matrix V given in (2.14).

From the expression (4.3), a bivariate elliptical t generator is written using FORTRAN IV with single precision and includes three subroutines. They are (1) uniform generator which is used in generating normal variates, (2) bivariate normal generator from which $\underline{X}$ is obtained using the Marsaglia's(1961) polar method, (3) Chi-square generator which simulates U as a sum of squares of n independent standard univariate normal variates.

In this thesis, V is given as $V = \frac{n-2}{n}\begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$, $n > 2$, $|\rho| < 1$, $V^{1/2}$ can be obtained by using the expression

$$V^{1/2} = \frac{1}{2} \frac{\sqrt{n-2}}{\sqrt{n}} \begin{bmatrix} \sqrt{1+\rho} + \sqrt{1-\rho} & \sqrt{1+\rho} - \sqrt{1-\rho} \\ \sqrt{1+\rho} - \sqrt{1-\rho} & \sqrt{1+\rho} + \sqrt{1-\rho} \end{bmatrix}. \tag{4.4}$$

The derivation of this expression is given in the Appendix D. Finally, the second component of the simulated random vector $\underline{T}$ is transformed to Z according to the thresholds.

§4.2.2 Contaminated normal distribution

Since the contaminated normal distribution is a linear combination of two bivariate normal distribution, namely, $N_2(\underline{0},V)$ and $N_2(\underline{0},\sigma^2V)$, where $V = \frac{1}{C}\begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$ and $C = (1-\epsilon) + \sigma^2\epsilon$. By the method given by Marsaglia(1961), this contaminated normal variates can be simulated by generating random variates from $N_2(\underline{0},V)$ with probability $(1-\epsilon)$ and from $N_2(\underline{0},\sigma^2V)$ with probability ϵ. Similar to the case of the bivariate t generator, the contaminated normal generator is written in FORTRAN IV with single precision, and includes Tausworthe uniform generator from U(0,1) and standard bivariate normal generator, but no chi-square generator is needed. The method of sample simulation is that by using the bivariate normal generator, a random vector $\underline{X}$ is simulated from the distribution $N_2(\underline{0},I_2)$, and an independent uniform random number is generated by using the uniform generator. If this random number is strictly greater than ϵ, then $\underline{X}$ is transformed into $\underline{Y}$ with $\underline{Y} = V^{1/2}\cdot\underline{X}$, where $V^{1/2}$ can be shown to be

$$V^{1/2} = \frac{1}{2\sqrt{C}}\begin{bmatrix} \sqrt{1+\rho}+\sqrt{1-\rho} & \sqrt{1+\rho}-\sqrt{1-\rho} \\ \sqrt{1+\rho}-\sqrt{1-\rho} & \sqrt{1+\rho}+\sqrt{1-\rho} \end{bmatrix}. \quad \text{--- (4.5)}$$

The derivation of the expression is similar to that of (4.4). Otherwise, if this random number is less than or equal to ϵ, then $\underline{X}$

is transformed into $\underline{T}$ with $\underline{T} = \sigma \cdot V^{1/2} \cdot \underline{X}$. Obviously, in the first case, $\underline{T}$ is distributed as $N_2(\underline{0}, V)$ and the probability of its occurance is $(1-\epsilon)$. Similarly, in the second case, $\underline{T}$ is distributed as $N_2(0, \sigma^2 \cdot V)$ with the probability of occurance ϵ. Thus, the simulated random vector $\underline{T}$ is our desired contaminated normal variate according to the method of Marsaglia(1961). On the same way, the second component of $\underline{T}$ will be further transformed into the discrete random variate Z based on the thresholds α_2 and α_3.

Chapter 5 Findings and conclusion

§5.1 Findings

The results of the simulation studies are summarized in the table 1.1 to the table 4.12(see the table section of this thesis). During the estimation procedure, it was observed that, in most cases, the estimates converge smoothly. About four percentage of estimations are considered to be diverged. In addition, in most cases, the value of the objective function and the gradient vector decrease rapidly in the first few iterations, and the objective function can attain its minimum in about 25 iterations. From the results obtained, we have some findings for two members of elliptical distribution family, and they are presented in Sections 5.1.1 and 5.1.2 respectively.

§5.1.1 Elliptical t distribution

Firstly, in most cases, when the sample size is large, the means of the estimates obtained under the elliptical t distribution assumption are near the true values of the parameters α_2, α_3 and ρ. However, large discrepancies between the means of the estimates and the true values are observed in the cases where the sample size is small(e.g. N =20). Moreover, By using the normal theory MLE, the discrepancies also occur not only in the situation where the sample size is small, but also in the situation where the degree of freedom of the true elliptical t distribution is low(e.g. d.f. = 5). This means that the large discrepancies between the estimates and the true values of parameters occur when the deviation of the sample distribution from normality is large or the parameters are estimated with small sample size.

Secondly, in the simulation studies, we noted that the magnitudes of most root mean squares obtained under the elliptical t distribution assumption and those obtained under the incorrect normal distribution assumption range from 0.1 to 0.3. As we expect, the root mean square tends to be smaller in magnitude when the sample size is large. The reason is that the maximum likelihood estimator is a consistent estimator, and the root mean square is an alternative measure of the sample standard error of the estimate, so it is expected to decrease if the sample size increase. Nevertheless, we noted that most root mean squares of the polyserial correlation estimates under the elliptical t distribution assumption are greater than that under the normality assumption, it is believed that the differences may be due to random errors.

5.1.2 Contaminated normal distribution

Firstly, in most cases, under the contaminated normal distribution assumption, the means of the estimates obtained with large sample size(i.e. N=100) are near the true values of the parameters to be estimated, but the differences between the means of the estimates and the true values are also noticeable when the sample size is small. Moreover, under the incorrect normal distribution assumption, larger discrepancies between the means of the estimates and their corresponding true values are observed. It is especially true in the cases where the sample size is small(e.g. N=20) and the deviation of the sample distribution from normality is large(e.g. ϵ=0.5 and σ=0.1). This phenomenon is prominent in the estimates of the thresholds.

Secondly, with larger sample size, smaller root mean squares are also obtained. For example, when the sample size is 100, most root mean squares of the estimates under the contaminated normal assumption are less than 0.2. That is, lower bias can also be obtained if the sample size is increased.

§5.2 Conclusion

In this thesis, we consider the estimation of the polyserial correlation on two familiar members of the bivariate elliptical distribution family, namely, elliptical t and contaminated normal distribution. A method for obtaining the estimates of the polyserial correlation and thresholds using the maximum likelihood method is established.

Based on the findings from the simulation studies, we can see that, under the two elliptical distribution assumptions, the estimates obtained by using maximum likelihood method with large sample size are not far away from the true values, and the root means squares of estimates tend to decrease if a large sample size is used.

On the other hand, from the results, we are convinced that the estimators are not robust against normality assumption. This means that the estimates obtained under the elliptical distribution assumptions and those obtained under the incorrect normality assumption are different. The differences are large especially in the situations where the deviation of the sample distribution from normality assumption is large. Thus, only if the sample distribution do not deviate much from normality, the estimates obtained under the normality assumption can be used as a crude estimates for the true values of the parameters, and special attention should be paid to the interpretation of the results.

The results obtained in this study are only based on the bivariate elliptical distribution family, no conclusion about the behaviour of the estimates in higher dimensional space can be made. As an extension of the research, a polyserial correlation model, in which X is an observable random vector and has a joint multivariate elliptical distribution with another latent variable Y, can be developed, and the maximum likelihood method can also be used for estimating the parameters of the model.

Table 1.1
Means of Estimates
(t distribution vs normal distribution)
True α_2 = -1.0
True α_3 = 0.0

True ρ	Sample size	Par.	d.f. = 5 t-Dist.	d.f. = 5 normal	d.f. = 10 t-Dist.	d.f. = 10 normal
0.0	20	$\alpha(2)$	-0.9938	-1.1576	-1.1909	-1.1828
		$\alpha(3)$	-0.0306	-0.0319	0.0153	-0.0139
		ρ	-0.0177	0.0986	0.0355	0.0524
	40	$\alpha(2)$	-1.0073	-1.1736	-0.9962	-1.0633
		$\alpha(3)$	0.0039	0.0023	0.0103	0.0263
		ρ	-0.0134	0.0441	0.0285	0.0162
	100	$\alpha(2)$	-1.0437	-1.1675	-1.0215	-1.0729
		$\alpha(3)$	-0.0246	-0.0171	-0.0054	0.0136
		ρ	0.0511	0.0315	0.0346	0.0270
0.2	20	$\alpha(2)$	-0.9583	-1.1814	-1.1593	-1.2180
		$\alpha(3)$	0.0064	-0.0330	0.2166	0.2389
		ρ	0.2364	0.2938	0.2166	0.2389
	40	$\alpha(2)$	-0.9815	-1.1557	-1.0052	-1.0741
		$\alpha(3)$	0.0171	-0.0245	0.0122	0.0225
		ρ	0.2295	0.2512	0.2012	0.1828
	100	$\alpha(2)$	-1.0367	-1.1631	-1.0367	-1.0801
		$\alpha(3)$	-0.0330	-0.0282	-0.0049	0.0098
		ρ	0.2441	0.2281	0.2302	0.2115
0.5	20	$\alpha(2)$	-1.0163	-1.2337	-1.0524	-1.1694
		$\alpha(3)$	-0.0086	-0.0156	0.0150	0.0104
		ρ	0.5167	0.5596	0.5202	0.5506
	40	$\alpha(2)$	-1.0159	-1.1538	-1.0195	-1.0776
		$\alpha(3)$	0.0012	-0.0188	0.0184	0.0245
		ρ	0.5429	0.5428	0.5229	0.5069
	100	$\alpha(2)$	-1.0269	-1.1453	-1.0287	-1.0671
		$\alpha(3)$	-0.0176	-0.0080	-0.0027	0.0140
		ρ	0.5352	0.5257	0.5313	0.5128
0.7	20	$\alpha(2)$	-1.0835	-1.3215	-1.1028	-1.0614
		$\alpha(3)$	0.0181	0.0029	0.0387	0.0036
		ρ	0.6795	0.7228	0.6623	0.7310
	40	$\alpha(2)$	-1.0543	-1.1164	-1.0443	-1.0587
		$\alpha(3)$	0.0129	-0.0139	0.0136	0.0198
		ρ	0.7075	0.7268	0.7102	0.7165
	100	$\alpha(2)$	-1.0396	-1.1158	-1.0208	-1.0515
		$\alpha(3)$	-0.0046	-0.0071	-0.0004	0.0187
		ρ	0.7389	0.7268	0.7115	0.7072

't-dist.' denotes estimates based on the t distribution assumption.
'normal' denotes estimates based on the normality assumption.

Table 1.2
Means of Estimates
(t distribution vs normal distribution)
True α_2 = -0.5
True α_3 = 0.5

True ρ	Sample size	Par.	d.f. = 5 t-Dist.	d.f. = 5 normal	d.f. = 10 t-Dist.	d.f. = 10 normal
0.0	20	$\alpha(2)$	-0.5487	-0.6163	-0.5709	-0.5744
		$\alpha(3)$	0.5136	0.6339	0.5196	0.5306
		ρ	0.0401	0.1208	0.0587	0.0453
	40	$\alpha(2)$	-0.5450	-0.6205	-0.5741	-0.5458
		$\alpha(3)$	0.5034	0.6540	0.4347	0.5778
		ρ	0.0087	0.0396	0.0404	0.0204
	100	$\alpha(2)$	-0.5257	-0.6055	-0.5428	-0.5374
		$\alpha(3)$	0.4796	0.6296	0.4813	0.5565
		ρ	0.0277	0.0183	0.0136	0.0242
0.2	20	$\alpha(2)$	-0.5359	-0.6080	-0.5762	-0.5639
		$\alpha(3)$	0.5204	0.6271	0.4727	0.5927
		ρ	0.2545	0.3174	0.2238	0.2621
	40	$\alpha(2)$	-0.5370	-0.6185	-0.5424	-0.5384
		$\alpha(3)$	0.5403	0.6472	0.4936	0.5735
		ρ	0.2029	0.2405	0.2325	0.2282
	100	$\alpha(2)$	-0.5256	-0.6115	-0.5228	-0.5349
		$\alpha(3)$	0.4305	0.6192	0.4957	0.5618
		ρ	0.2324	0.2262	0.2102	0.2258
0.5	20	$\alpha(2)$	-0.4959	-0.5666	-0.5434	-0.5381
		$\alpha(3)$	0.4916	0.6571	0.5184	0.5806
		ρ	0.4927	0.5702	0.4448	0.5661
	40	$\alpha(2)$	-0.5416	-0.6001	-0.5467	-0.5443
		$\alpha(3)$	0.5100	0.6078	0.4961	0.5798
		ρ	0.4636	0.5340	0.5229	0.5233
	100	$\alpha(2)$	-0.5356	-0.6008	-0.5230	-0.5300
		$\alpha(3)$	0.4880	0.5956	0.5060	0.5624
		ρ	0.5554	0.5231	0.5109	0.5182
0.7	20	$\alpha(2)$	-0.5410	-0.5853	-0.5124	-0.5453
		$\alpha(3)$	0.5316	0.5802	0.5035	0.5678
		ρ	0.6359	0.7221	0.6511	0.7556
	40	$\alpha(2)$	-0.5406	-0.5954	-0.5287	-0.5207
		$\alpha(3)$	0.5193	0.5777	0.5221	0.5650
		ρ	0.6369	0.7290	0.6939	0.7321
	100	$\alpha(2)$	-0.5349	-0.5825	-0.5129	-0.5152
		$\alpha(3)$	0.4902	0.5725	0.5056	0.5486
		ρ	0.7322	0.7259	0.7175	0.7218

't-dist.' denotes estimates based on the t distribution assumption.
'normal' denotes estimates based on the normality assumption.

Table 1.3

Means of Estimates

(t distribution vs normal distribution)

True α_2 = -0.3

True α_3 = 0.7

True ρ	Sample size	Par.	d.f. = 5 t-Dist.	d.f. = 5 normal	d.f. = 10 t-Dist.	d.f. = 10 normal
0.0	20	$\alpha(2)$	-0.3670	-0.4043	-0.3909	-0.3332
		$\alpha(3)$	0.6724	0.8670	0.6594	0.8046
		ρ	0.0025	0.1269	0.0195	0.0617
	40	$\alpha(2)$	-0.3333	-0.4069	-0.3294	-0.3073
		$\alpha(3)$	0.5946	0.8592	0.6419	0.7741
		ρ	0.0452	0.0584	0.0455	0.0365
	100	$\alpha(2)$	-0.2842	-0.3782	-0.2961	-0.3177
		$\alpha(3)$	0.5586	0.8360	0.6103	0.7704
		ρ	0.0632	0.0223	0.0249	0.0296
0.2	20	$\alpha(2)$	-0.3811	-0.4046	-0.3704	-0.3324
		$\alpha(3)$	0.5798	0.8455	0.6709	0.8384
		ρ	0.1651	0.3066	0.1924	0.2763
	40	$\alpha(2)$	-0.3199	-0.4037	-0.3498	-0.3068
		$\alpha(3)$	0.5800	0.8549	0.6208	0.7901
		ρ	0.1986	0.2432	0.2316	0.2348
	100	$\alpha(2)$	-0.2791	-0.3763	-0.3043	-0.3088
		$\alpha(3)$	0.5691	0.8406	0.6234	0.7756
		ρ	0.2053	0.2190	0.2030	0.2290
0.5	20	$\alpha(2)$	-0.3549	-0.3700	-0.3777	-0.3653
		$\alpha(3)$	0.6051	0.8780	0.6404	0.8275
		ρ	0.4166	0.5498	0.3830	0.5715
	40	$\alpha(2)$	-0.3052	-0.3707	-0.3466	-0.3221
		$\alpha(3)$	0.6071	0.8241	0.6430	0.7754
		ρ	0.4334	0.5323	0.4725	0.5274
	100	$\alpha(2)$	-0.2821	-0.3599	-0.3144	-0.3125
		$\alpha(3)$	0.5719	0.8191	0.6338	0.7629
		ρ	0.4954	0.5285	0.5061	0.5178
0.7	20	$\alpha(2)$	-0.3372	-0.3833	-0.3681	-0.3461
		$\alpha(3)$	0.6690	0.8252	0.6819	0.7656
		ρ	0.6018	0.7336	0.5513	0.7534
	40	$\alpha(2)$	-0.2937	-0.3627	-0.3605	-0.3207
		$\alpha(3)$	0.6504	0.7866	0.6689	0.7592
		ρ	0.6279	0.7331	0.6434	0.7334
	100	$\alpha(2)$	-0.2949	-0.3569	-0.3267	-0.3066
		$\alpha(3)$	0.6167	0.7972	0.6659	0.7527
		ρ	0.6853	0.7229	0.6748	0.7172

't-dist.' denotes estimates based on the t distribution assumption.
'normal' denotes estimates based on the normality assumption.

Table 2.1
Means of Estimates
(contaminated normal vs normal distribution)

True α_2 = -1.0

True α_3 = 0.0

True Corr. = 0.0

N	ϵ	Para.	$\sigma = 0.1$		$\sigma = 0.5$	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	-1.0403	-1.2345	-1.1272	-1.2206
		$\alpha(3)$	0.0402	0.0501	0.0505	0.0521
		ρ	-0.0417	-0.0218	-0.0401	-0.0255
	0.3	$\alpha(2)$	-1.2349	-1.3380	-1.1643	-1.5231
		$\alpha(3)$	0.0041	0.0686	0.0423	0.0558
		ρ	-0.0121	-0.0297	-0.0292	-0.0216
	0.5	$\alpha(2)$	-1.2019	-1.6688	-1.2004	-1.6279
		$\alpha(3)$	-0.0017	0.0828	0.0446	0.0545
		ρ	-0.0037	0.0682	-0.0354	-0.0157
40	0.1	$\alpha(2)$	-1.0248	-1.1360	-1.0677	-1.1243
		$\alpha(3)$	0.0335	0.0153	0.0175	0.0163
		ρ	-0.0406	-0.0404	-0.0447	-0.0448
	0.3	$\alpha(2)$	-1.1932	-1.3300	-1.1276	-1.2947
		$\alpha(3)$	0.0092	0.0148	0.0167	0.0165
		ρ	-0.0358	-0.0277	-0.0374	-0.0387
	0.5	$\alpha(2)$	-1.0736	-1.5233	-1.1073	-1.4254
		$\alpha(3)$	0.0004	0.0153	0.0118	0.0162
		ρ	-0.0350	0.0022	-0.0347	-0.0292
100	0.1	$\alpha(2)$	-1.0144	-1.0936	-1.0205	-1.0780
		$\alpha(3)$	0.0058	0.0023	0.0029	0.0026
		ρ	-0.0214	-0.0200	-0.0209	-0.0215
	0.3	$\alpha(2)$	-1.0243	-1.2519	-1.0280	-1.2107
		$\alpha(3)$	0.0014	0.0021	0.0026	0.0027
		ρ	-0.0183	-0.0082	-0.0160	-0.0161
	0.5	$\alpha(2)$	-1.0070	-1.4071	-1.0053	-1.3316
		$\alpha(3)$	-0.0064	0.0020	0.0011	0.0029
		ρ	-0.0158	0.0152	-0.0146	-0.0106

'Cont.-N' denotes the estimates based on the contaminated normal assumption.
'Normal' denotes the estimates based on the normality assumption.

Table 2.2
Means of Estimates
(contaminated normal vs normal distribution)

True α_2 = -1.0

True α_3 = 0.0

True Corr. = 0.2

N	ϵ	Para.	$\sigma = 0.1$		$\sigma = 0.5$	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	-1.1738	-1.2300	-1.1796	-1.2179
		$\alpha(3)$	0.0311	0.0582	0.0533	0.0572
		ρ	0.1590	0.1773	0.1682	0.1717
	0.3	$\alpha(2)$	-1.2949	-1.3532	-1.2027	-1.4474
		$\alpha(3)$	-0.0064	0.0621	0.0465	0.0582
		ρ	0.1928	0.1946	0.1813	0.1929
	0.5	$\alpha(2)$	-1.2902	-1.6881	-1.2165	-1.8813
		$\alpha(3)$	0.0009	0.0737	0.0379	0.0569
		ρ	0.2024	0.2483	0.1784	0.2161
40	0.1	$\alpha(2)$	-1.0202	-1.1254	-1.0588	-1.1150
		$\alpha(3)$	0.0340	0.0233	0.0245	0.0237
		ρ	0.1623	0.1666	0.1549	0.1614
	0.3	$\alpha(2)$	-1.0556	-1.2820	-1.0840	-1.2538
		$\alpha(3)$	0.0197	0.0222	0.0219	0.0231
		ρ	0.1700	0.1933	0.1641	0.1845
	0.5	$\alpha(2)$	-1.0506	-1.4006	-1.0805	-1.3743
		$\alpha(3)$	0.0029	0.0233	0.0163	0.0237
		ρ	0.1656	0.2322	0.1676	0.2133
100	0.1	$\alpha(2)$	-1.0185	-1.0924	-1.0236	-1.0795
		$\alpha(3)$	0.0043	0.0049	0.0050	0.0048
		ρ	0.1796	0.1836	0.1771	0.1828
	0.3	$\alpha(2)$	-1.0272	-1.2432	-1.0337	-1.2096
		$\alpha(3)$	0.0030	0.0051	0.0045	0.0046
		ρ	0.1855	0.2122	0.1841	0.2065
	0.5	$\alpha(2)$	-1.0021	-1.3798	-1.0135	-1.3235
		$\alpha(3)$	0.0154	0.0065	0.0030	0.0052
		ρ	0.1928	0.2579	0.1887	0.2376

'Cont.-N' denotes the estimates based on the contaminated normal assumption.
'Normal' denotes the estimates based on the normality assumption.

Table 2.3
Means of Estimates
(contaminated normal vs normal distribution)

True α_2 = -1.0

True α_3 = 0.0

True Corr. = 0.5

N	ϵ	Para.	σ = 0.1		σ = 0.5	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	-1.1943	-1.2644	-1.2203	-1.2568
		$\alpha(3)$	0.0327	0.0716	0.0612	0.0641
		ρ	0.5187	0.5218	0.5058	0.5182
	0.3	$\alpha(2)$	-1.2794	-1.4762	-1.3634	-1.3974
		$\alpha(3)$	0.0022	0.0686	0.0480	0.0594
		ρ	0.5160	0.5302	0.5118	0.5455
	0.5	$\alpha(2)$	-1.4506	-1.4965	-1.3653	-1.4777
		$\alpha(3)$	0.0039	0.0821	0.0402	0.0606
		ρ	0.5241	0.5393	0.5049	0.5664
40	0.1	$\alpha(2)$	-1.0640	-1.1224	-1.0773	-1.1168
		$\alpha(3)$	0.0176	0.0369	0.0333	0.0341
		ρ	0.4847	0.4994	0.4845	0.4935
	0.3	$\alpha(2)$	-1.1965	-1.2715	-1.1175	-1.2568
		$\alpha(3)$	0.0109	0.0278	0.0265	0.0294
		ρ	0.4836	0.5259	0.4890	0.5295
	0.5	$\alpha(2)$	-1.1043	-1.3397	-1.1483	-1.3458
		$\alpha(3)$	0.0046	0.0417	0.0230	0.0312
		ρ	0.4786	0.5424	0.4840	0.5638
100	0.1	$\alpha(2)$	-1.0309	-1.0811	-1.0291	-1.0739
		$\alpha(3)$	0.0051	0.0194	0.0158	0.0175
		ρ	0.4987	0.5115	0.4972	0.5109
	0.3	$\alpha(2)$	-1.0326	-1.1859	-1.0374	-1.1727
		$\alpha(3)$	0.0048	0.0199	0.0130	0.0171
		ρ	0.5021	0.5491	0.5014	0.5481
	0.5	$\alpha(2)$	-1.0115	-1.2659	-1.0158	-1.2457
		$\alpha(3)$	0.0109	0.0218	0.0109	0.0171
		ρ	0.5011	0.5985	0.5019	0.5898

'Cont.-N' denotes the estimates based on the contaminated normal assumption.
'Normal' denotes the estimates based on the normality assumption.

Table 2.4
Means of Estimates
(contaminated normal vs normal distribution)

True α_2 = -1.0
True α_3 = 0.0
True Corr. = 0.7

N	ϵ	Para.	σ = 0.1		σ = 0.5	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	-1.2199	-1.1471	-1.1638	-1.1766
		$\alpha(3)$	0.0414	0.0705	0.0479	0.0510
		ρ	0.7241	0.7253	0.7163	0.7298
	0.3	$\alpha(2)$	-1.3006	-1.2237	-1.1769	-1.2454
		$\alpha(3)$	0.0067	0.0631	0.0454	0.0576
		ρ	0.7249	0.7316	0.7072	0.7327
	0.5	$\alpha(2)$	-1.2397	-1.6509	-1.1979	-1.3063
		$\alpha(3)$	0.0082	0.0752	0.0354	0.0609
		ρ	0.7185	0.7152	0.7000	0.7520
40	0.1	$\alpha(2)$	-1.0354	-1.0735	-1.0481	-1.0793
		$\alpha(3)$	0.0245	0.0422	0.0365	0.0385
		ρ	0.6975	0.7120	0.6977	0.7118
	0.3	$\alpha(2)$	-1.0463	-1.1462	-1.0654	-1.1454
		$\alpha(3)$	0.0128	0.0347	0.0297	0.0403
		ρ	0.6927	0.7382	0.6978	0.7387
	0.5	$\alpha(2)$	-1.0232	-1.1816	-1.0553	-1.1382
		$\alpha(3)$	0.0055	0.0427	0.0264	0.0373
		ρ	0.6898	0.7516	0.6908	0.7629
100	0.1	$\alpha(2)$	-1.0218	-1.0551	-1.0229	-1.0552
		$\alpha(3)$	0.0030	0.0128	0.0092	0.0098
		ρ	0.6999	0.7129	0.6985	0.7123
	0.3	$\alpha(2)$	-1.0166	-1.1099	-1.0195	-1.1133
		$\alpha(3)$	0.0041	0.0132	0.0083	0.0089
		ρ	0.6999	0.7448	0.7003	0.7450
	0.5	$\alpha(2)$	-1.0012	-1.1492	-1.0062	-1.1591
		$\alpha(3)$	0.0032	0.0161	0.0070	0.0094
		ρ	0.6960	0.7795	0.6958	0.7758

'Cont.-N' denotes the estimates based on the contaminated normal assumption.
'Normal' denotes the estimates based on the normality assumption.

Table 2.5
Means of Estimates
(contaminated normal vs normal distribution)

True α_2 = -0.5
True α_3 = 0.5
True Corr. = 0.0

N	ϵ	Para.	σ = 0.1		σ = 0.5	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	-0.4809	-0.5681	-0.4820	-0.5187
		$\alpha(3)$	0.6362	0.7389	0.6491	0.6915
		ρ	0.0153	0.0159	0.0138	0.0135
	0.3	$\alpha(2)$	-0.5137	-0.7855	-0.5073	-0.6221
		$\alpha(3)$	0.6447	0.9317	0.6544	0.7925
		ρ	0.0096	0.0373	0.0307	0.0294
	0.5	$\alpha(2)$	-0.5049	-0.9208	-0.5066	-0.6779
		$\alpha(3)$	0.8356	1.2569	0.6215	0.8539
		ρ	0.0331	-0.0121	0.0196	0.0195
40	0.1	$\alpha(2)$	-0.5168	-0.6173	-0.5268	-0.5709
		$\alpha(3)$	0.5467	0.6480	0.5713	0.6136
		ρ	-0.0049	-0.0074	-0.0110	-0.0122
	0.3	$\alpha(2)$	-0.5274	-0.8162	-0.5357	-0.6677
		$\alpha(3)$	0.5484	0.8473	0.5682	0.7030
		ρ	-0.0083	0.0074	-0.0062	-0.0084
	0.5	$\alpha(2)$	-0.5152	-0.9918	-0.5310	-0.7498
		$\alpha(3)$	0.6125	1.0843	0.5576	0.7913
		ρ	0.0018	0.0105	-0.0041	0.0036
100	0.1	$\alpha(2)$	-0.5118	-0.6077	-0.5158	-0.5594
		$\alpha(3)$	0.5044	0.6012	0.5147	0.5580
		ρ	-0.0029	-0.0035	-0.0054	-0.0062
	0.3	$\alpha(2)$	-0.5145	-0.7999	-0.5142	-0.6469
		$\alpha(3)$	0.4998	0.7939	0.5172	0.6515
		ρ	-0.0014	0.0000	-0.0017	-0.0042
	0.5	$\alpha(2)$	-0.4981	-1.0021	-0.5141	-0.7400
		$\alpha(3)$	0.5233	1.0321	0.5199	0.7527
		ρ	0.0169	0.0195	0.0017	0.0024

'Cont.-N' denotes the estimates based on the contaminated normal assumption.
'Normal' denotes the estimates based on the normality assumption.

Table 2.6
Means of Estimates
(contaminated normal vs normal distribution)

True α_2 = -0.5

True α_3 = 0.5

True Corr. = 0.2

N	ϵ	Para.	σ = 0.1		σ = 0.5	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	-0.4753	-0.5542	-0.4690	-0.5040
		$\alpha(3)$	0.6175	0.7297	0.6448	0.6869
		ρ	0.2140	0.2292	0.2224	0.2284
	0.3	$\alpha(2)$	-0.4858	-0.7534	-0.4842	-0.5914
		$\alpha(3)$	0.6914	0.9832	0.6461	0.7800
		ρ	0.2343	0.2611	0.2369	0.2546
	0.5	$\alpha(2)$	-0.4892	-0.9036	-0.4896	-0.6718
		$\alpha(3)$	0.6593	1.1783	0.6263	0.8524
		ρ	0.2194	0.2235	0.2224	0.2566
40	0.1	$\alpha(2)$	-0.4889	-0.5939	-0.5038	-0.5456
		$\alpha(3)$	0.5445	0.6421	0.5717	0.6136
		ρ	0.1382	0.1984	0.1356	0.1917
	0.3	$\alpha(2)$	-0.5018	-0.7871	-0.5122	-0.6362
		$\alpha(3)$	0.5329	0.8237	0.5671	0.6983
		ρ	0.1975	0.2247	0.1357	0.2045
	0.5	$\alpha(2)$	-0.4944	-0.9595	-0.5138	-0.7180
		$\alpha(3)$	0.5943	1.0458	0.5545	0.7814
		ρ	0.1391	0.2536	0.1388	0.2229
100	0.1	$\alpha(2)$	-0.5108	-0.6060	-0.5142	-0.5574
		$\alpha(3)$	0.4925	0.5890	0.5060	0.5472
		ρ	0.1927	0.2043	0.1925	0.1986
	0.3	$\alpha(2)$	-0.5068	-0.7886	-0.5083	-0.6371
		$\alpha(3)$	0.4349	0.7719	0.5092	0.6367
		ρ	0.1987	0.2364	0.1972	0.2187
	0.5	$\alpha(2)$	-0.4759	-0.9705	-0.5059	-0.7193
		$\alpha(3)$	0.5188	0.9990	0.5165	0.7380
		ρ	0.2192	0.2932	0.2047	0.2517

'Cont.-N' denotes the estimates based on the contaminated normal assumption.
'Normal' denotes the estimates based on the normality assumption.

Table 2.7
Means of Estimates
(contaminated normal vs normal distribution)

True α_2 = -0.5
True α_3 = 0.5
True Corr. = 0.5

N	ϵ	Para.	$\sigma = 0.1$ Cont.-N.	$\sigma = 0.1$ Normal	$\sigma = 0.5$ Cont.-N.	$\sigma = 0.5$ Normal
20	0.1	$\alpha(2)$	-0.4755	-0.5404	-0.4677	-0.4993
		$\alpha(3)$	0.5824	0.6852	0.6052	0.6386
		ρ	0.5272	0.5401	0.5294	0.5434
	0.3	$\alpha(3)$	0.6226	0.8532	0.6155	0.7152
		ρ	0.5216	0.5798	0.5441	0.5670
	0.5	$\alpha(2)$	-0.4776	-0.8219	-0.4726	-0.6123
		$\alpha(3)$	0.7296	1.0337	0.6044	0.7830
		ρ	0.4710	0.6014	0.5437	0.5918
40	0.1	$\alpha(2)$	-0.4849	-0.5796	-0.4939	-0.5302
		$\alpha(3)$	0.5284	0.6195	0.5541	0.5889
		ρ	0.4955	0.5209	0.4971	0.5114
	0.3	$\alpha(2)$	-0.5028	-0.7413	-0.5099	-0.6129
		$\alpha(3)$	0.5266	0.7706	0.5533	0.6583
		ρ	0.5036	0.5702	0.5034	0.5480
	0.5	$\alpha(2)$	-0.4844	-0.8588	-0.5004	-0.6580
		$\alpha(3)$	0.5886	0.9384	0.5511	0.7248
		ρ	0.5074	0.6135	0.5144	0.5903
100	0.1	$\alpha(2)$	-0.5020	-0.5859	-0.5016	-0.5383
		$\alpha(3)$	0.4903	0.5752	0.5017	0.5366
		ρ	0.4989	0.5237	0.4998	0.5138
	0.3	$\alpha(2)$	-0.4986	-0.7348	-0.5020	-0.6074
		$\alpha(3)$	0.4839	0.7209	0.5018	0.6034
		ρ	0.5096	0.5828	0.5073	0.5543
	0.5	$\alpha(2)$	-0.4584	-0.8551	-0.4971	-0.6612
		$\alpha(3)$	0.5067	0.8809	0.5080	0.6776
		ρ	0.5167	0.6401	0.5085	0.5943

'Cont.-N' denotes the estimates based on the contaminated normal assumption.
'Normal' denotes the estimates based on the normality assumption.

Table 2.8

Means of Estimates

(contaminated normal vs normal distribution)

True α_2 = -0.5

True α_3 = 0.5

True Corr. = 0.7

N	ϵ	Para.	$\sigma = 0.1$		$\sigma = 0.5$	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	-0.4365	-0.5116	-0.4551	-0.4794
		$\alpha(3)$	0.5786	0.6472	0.5842	0.6095
		ρ	0.7237	0.7430	0.7316	0.7422
	0.3	$\alpha(2)$	-0.4619	-0.6043	-0.4742	-0.5333
		$\alpha(3)$	0.6585	0.7339	0.5905	0.6682
		ρ	0.7258	0.7659	0.7286	0.7586
	0.5	$\alpha(2)$	-0.4660	-0.6803	-0.4625	-0.5501
		$\alpha(3)$	0.6437	0.7984	0.5708	0.6742
		ρ	0.7203	0.7616	0.7345	0.7825
40	0.1	$\alpha(2)$	-0.4666	-0.5405	-0.4765	-0.5030
		$\alpha(3)$	0.5260	0.5994	0.5490	0.5773
		ρ	0.7123	0.7362	0.7162	0.7300
	0.3	$\alpha(2)$	-0.4827	-0.6634	-0.4954	-0.5684
		$\alpha(3)$	0.5208	0.6997	0.5396	0.6139
		ρ	0.7155	0.7794	0.7171	0.7599
	0.5	$\alpha(2)$	-0.4627	-0.7426	-0.4922	-0.5868
		$\alpha(3)$	0.5505	0.8091	0.5356	0.6504
		ρ	0.7140	0.8058	0.7280	0.7950
100	0.1	$\alpha(2)$	-0.4897	-0.5578	-0.4907	-0.5189
		$\alpha(3)$	0.4941	0.5619	0.5020	0.5297
		ρ	0.7052	0.7299	0.7069	0.7212
	0.3	$\alpha(2)$	-0.4901	-0.6696	-0.4966	-0.5747
		$\alpha(3)$	0.4763	0.6612	0.4968	0.5703
		ρ	0.7115	0.7787	0.7100	0.7544
	0.5	$\alpha(2)$	-0.4538	-0.7407	-0.4909	-0.6038
		$\alpha(3)$	0.4810	0.7565	0.4997	0.6128
		ρ	0.7173	0.8190	0.7122	0.7386

'Cont.-N' denotes the estimates based on the contaminated normal assumption.
'Normal' denotes the estimates based on the normality assumption.

Table 2.9
Means of Estimates
(contaminated normal vs normal distribution)

True α_2 = -0.3
True α_3 = 0.7
True Corr. = 0.0

N	ϵ	Para.	σ = 0.1		σ = 0.5	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	-0.2815	-0.3455	-0.2358	-0.2519
		$\alpha(3)$	0.8197	0.9320	0.8577	0.9074
		ρ	0.0022	0.0040	0.0075	0.0073
	0.3	$\alpha(2)$	-0.3601	-0.5790	-0.2597	-0.3152
		$\alpha(3)$	0.8737	1.2183	0.8755	1.0326
		ρ	0.0228	-0.0025	0.0196	0.0216
	0.5	$\alpha(2)$	-0.3805	-0.7766	-0.2656	-0.3647
		$\alpha(3)$	1.2393	1.6148	0.9181	1.1754
		ρ	-0.0002	0.0072	0.0121	0.0119
40	0.1	$\alpha(2)$	-0.2942	-0.3896	-0.2756	-0.3001
		$\alpha(3)$	0.7555	0.8436	0.7738	0.8245
		ρ	-0.0056	-0.0079	-0.0109	-0.0111
	0.3	$\alpha(2)$	-0.3387	-0.6069	-0.2796	-0.3561
		$\alpha(3)$	0.6812	1.0048	0.7738	0.9340
		ρ	0.0052	0.0012	-0.0134	-0.0132
	0.5	$\alpha(2)$	-0.3547	-0.8158	-0.2821	-0.4069
		$\alpha(3)$	0.7662	1.2262	0.7858	1.0731
		ρ	0.0070	0.0012	-0.0131	-0.0099
100	0.1	$\alpha(2)$	-0.2942	-0.3985	-0.2907	-0.3189
		$\alpha(3)$	0.6963	0.7866	0.7095	0.7611
		ρ	-0.0157	-0.0166	-0.0163	-0.0168
	0.3	$\alpha(2)$	-0.3081	-0.6202	-0.2975	-0.3832
		$\alpha(3)$	0.6774	0.9494	0.7050	0.8675
		ρ	-0.0090	-0.0088	-0.0084	-0.0104
	0.5	$\alpha(2)$	-0.3100	-0.8402	-0.2960	-0.4371
		$\alpha(3)$	0.6765	1.1681	0.7179	1.0079
		ρ	-0.0004	0.0145	0.0011	0.0034

'Cont.-N' denotes the estimates based on the contaminated normal assumption.
'Normal' denotes the estimates based on the normality assumption.

Table 2.10
Means of Estimates
(contaminated normal vs normal distribution)

True α_2 = -0.3
True α_3 = 0.7
True Corr. = 0.2

N	ϵ	Para.	σ = 0.1		σ = 0.5	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	-0.3177	-0.3625	-0.2419	-0.2613
		$\alpha(3)$	0.8109	0.9379	0.8681	0.9149
		ρ	0.1969	0.2078	0.2064	0.2107
	0.3	$\alpha(2)$	-0.3632	-0.5904	-0.2624	-0.3217
		$\alpha(3)$	0.9616	1.1805	0.8714	1.0215
		ρ	0.1902	0.2026	0.2022	0.2198
	0.5	$\alpha(2)$	-0.3818	-0.7708	-0.2618	-0.3596
		$\alpha(3)$	1.0108	1.4516	0.8917	1.1493
		ρ	0.1687	0.2522	0.2036	0.2329
40	0.1	$\alpha(2)$	-0.3102	-0.4031	-0.2878	-0.3137
		$\alpha(3)$	0.7454	0.8430	0.7782	0.8270
		ρ	0.1982	0.2080	0.1979	0.2041
	0.3	$\alpha(2)$	-0.3520	-0.6123	-0.2866	-0.3652
		$\alpha(3)$	0.6876	0.9938	0.7735	0.9280
		ρ	0.2035	0.2278	0.1887	0.2074
	0.5	$\alpha(2)$	-0.3633	-0.7974	-0.2875	-0.4090
		$\alpha(3)$	0.7625	1.1367	0.7853	1.0629
		ρ	0.1921	0.2492	0.1897	0.2311
100	0.1	$\alpha(2)$	-0.2964	-0.4016	-0.2923	-0.3204
		$\alpha(3)$	0.7039	0.7901	0.7156	0.7654
		ρ	0.1900	0.2025	0.1901	0.1966
	0.3	$\alpha(2)$	-0.3117	-0.6130	-0.2934	-0.3765
		$\alpha(3)$	0.6806	0.9446	0.7098	0.8674
		ρ	0.1967	0.2318	0.1888	0.2103
	0.5	$\alpha(2)$	-0.3112	-0.8145	-0.2931	-0.4270
		$\alpha(3)$	0.6780	1.1435	0.7205	0.9984
		ρ	0.2044	0.2789	0.1953	0.2425

'Cont.-N' denotes the estimates based on the contaminated normal assumption.
'Normal' denotes the estimates based on the normality assumption.

Table 2.11
Means of Estimates
(contaminated normal vs normal distribution)

True α_2 = -0.3
True α_3 = 0.7
True Corr. = 0.5

N	ϵ	Para.	σ = 0.1		σ = 0.5	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	-0.3004	-0.3667	-0.2617	-0.2815
		$\alpha(3)$	0.8241	0.8922	0.8384	0.8747
		ρ	0.5314	0.5386	0.5234	0.5380
	0.3	$\alpha(2)$	-0.3352	-0.5459	-0.2620	-0.3083
		$\alpha(3)$	0.9124	0.9919	0.8564	0.9757
		ρ	0.5112	0.5715	0.5171	0.5475
	0.5	$\alpha(2)$	-0.3726	-0.6734	-0.2685	-0.3409
		$\alpha(3)$	1.3408	1.2529	0.8714	1.1084
		ρ	0.4668	0.6428	0.5249	0.5775
40	0.1	$\alpha(2)$	-0.3073	-0.3927	-0.2949	-0.3174
		$\alpha(3)$	0.7457	0.8241	0.7689	0.8102
		ρ	0.5047	0.5255	0.5055	0.5191
	0.3	$\alpha(2)$	-0.3397	-0.5773	-0.2919	-0.3566
		$\alpha(3)$	0.7156	0.9394	0.7699	0.8964
		ρ	0.5124	0.5679	0.5027	0.5452
	0.5	$\alpha(2)$	-0.3686	-0.7323	-0.2971	-0.3961
		$\alpha(3)$	0.7764	1.1168	0.7687	0.9829
		ρ	0.4988	0.5928	0.4984	0.5772
100	0.1	$\alpha(2)$	-0.2952	-0.3935	-0.2958	-0.3199
		$\alpha(3)$	0.6921	0.7655	0.7060	0.7478
		ρ	0.5121	0.5357	0.5118	0.5260
	0.3	$\alpha(2)$	-0.3146	-0.5717	-0.2948	-0.3630
		$\alpha(3)$	0.6685	0.8759	0.6997	0.8239
		ρ	0.5190	0.5894	0.5093	0.5563
	0.5	$\alpha(2)$	-0.3099	-0.7238	-0.2967	-0.4030
		$\alpha(3)$	0.6657	1.0025	0.7038	0.9108
		ρ	0.5256	0.6431	0.5042	0.5933

'Cont.-N' denotes the estimates based on the contaminated normal assumption.
'Normal' denotes the estimates based on the normality assumption.

Table 2.12
Means of Estimates
(contaminated normal vs normal distribution)

True α_2 = -0.3
True α_3 = 0.7
True Corr. = 0.7

N	ϵ	Para.	σ = 0.1		σ = 0.5	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	-0.3094	-0.3773	-0.2895	-0.3055
		$\alpha(3)$	0.8137	0.8666	0.8267	0.8655
		ρ	0.7290	0.7489	0.7309	0.7419
	0.3	$\alpha(2)$	-0.3422	-0.5025	-0.2875	-0.3343
		$\alpha(3)$	1.4032	0.9682	0.8243	0.8923
		ρ	0.7252	0.7672	0.7321	0.7636
	0.5	$\alpha(2)$	-0.3722	-0.5772	-0.2911	-0.3464
		$\alpha(3)$	0.8172	1.0378	0.7989	0.9049
		ρ	0.7147	0.7730	0.7230	0.7596
40	0.1	$\alpha(2)$	-0.3091	-0.3911	-0.3061	-0.5030
		$\alpha(3)$	0.7486	0.7895	0.7508	0.5773
		ρ	0.7069	0.7329	0.7123	0.7300
	0.3	$\alpha(2)$	-0.3402	-0.5399	-0.2989	-0.3518
		$\alpha(3)$	0.7003	0.8507	0.7400	0.8250
		ρ	0.7180	0.7754	0.7125	0.7552
	0.5	$\alpha(2)$	-0.3557	-0.6436	-0.3071	-0.3829
		$\alpha(3)$	0.7113	0.9958	0.7732	0.9025
		ρ	0.7096	0.7882	0.7020	0.7785
100	0.1	$\alpha(2)$	-0.2971	-0.3818	-0.2959	-0.3152
		$\alpha(3)$	0.6945	0.7446	0.7016	0.7334
		ρ	0.7124	0.7331	0.7099	0.7239
	0.3	$\alpha(2)$	-0.3133	-0.5234	-0.2938	-0.3449
		$\alpha(3)$	0.6717	0.8108	0.6943	0.7821
		ρ	0.7164	0.7806	0.7111	0.7551
	0.5	$\alpha(2)$	-0.3137	-0.6347	-0.2977	-0.3739
		$\alpha(3)$	0.6658	0.9867	0.6987	0.8384
		ρ	0.7160	0.8132	0.7019	0.7808

'Cont.-N' denotes the estimates based on the contaminated normal assumption.
'Normal' denotes the estimates based on the normality assumption.

Table 3.1
RMS of Estimates
(t distribution vs normal distribution)

True α_2 = -1.0
True α_3 = 0.0

True ρ	Sample size	Par.	d.f. = 5		d.f. = 10	
			t-Dist.	normal	t-Dist.	normal
0.0	20	$\alpha(2)$	0.2798	0.3508	0.8684	0.8040
		$\alpha(3)$	0.2074	0.2417	0.2462	0.3502
		ρ	0.2984	0.2932	0.3342	0.2771
	40	$\alpha(2)$	0.2393	0.3159	0.2699	0.2881
		$\alpha(3)$	0.1462	0.1924	0.1787	0.2323
		ρ	0.2246	0.2118	0.2755	0.2417
	100	$\alpha(2)$	0.1858	0.2457	0.1562	0.1738
		$\alpha(3)$	0.0906	0.1063	0.1147	0.1293
		ρ	0.1811	0.1646	0.1630	0.1411
0.2	20	$\alpha(2)$	0.2458	0.5698	0.8846	0.7770
		$\alpha(3)$	0.2136	0.2456	0.2602	0.3408
		ρ	0.3148	0.2912	0.2633	0.2533
	40	$\alpha(2)$	0.2415	0.3097	0.2797	0.2925
		$\alpha(3)$	0.1527	0.2073	0.1812	0.2241
		ρ	0.1933	0.1995	0.2190	0.2253
	100	$\alpha(2)$	0.2229	0.2610	0.1535	0.1732
		$\alpha(3)$	0.0996	0.1106	0.1201	0.1253
		ρ	0.1804	0.1575	0.1693	0.1335
0.5	20	$\alpha(2)$	0.2034	0.6243	0.3546	0.7180
		$\alpha(3)$	0.2066	0.2499	0.2689	0.2984
		ρ	0.2449	0.2315	0.2372	0.2217
	40	$\alpha(2)$	0.2115	0.2994	0.2698	0.2770
		$\alpha(3)$	0.1792	0.2038	0.1984	0.2142
		ρ	0.1616	0.1708	0.1866	0.1608
	100	$\alpha(2)$	0.1531	0.2163	0.1390	0.1566
		$\alpha(3)$	0.1001	0.1138	0.1147	0.1157
		ρ	0.1322	0.1227	0.1184	0.0933
0.7	20	$\alpha(2)$	0.3073	0.9984	0.4847	0.3661
		$\alpha(3)$	0.1924	0.2256	0.2144	0.2575
		ρ	0.1917	0.1461	0.1730	0.1449
	40	$\alpha(2)$	0.2188	0.2423	0.2703	0.2853
		$\alpha(3)$	0.1663	0.1824	0.1892	0.2056
		ρ	0.1230	0.1200	0.1343	0.1195
	100	$\alpha(2)$	0.1534	0.1867	0.1428	0.1532
		$\alpha(3)$	0.0864	0.1000	0.1116	0.1081
		ρ	0.0826	0.0851	0.0705	0.0604

't-dist.' denotes estimates based on the t distribution assumption.
'normal' denotes estimates based on the normality assumption.

Table 3.2
RMS of Estimates
(t distribution vs normal distribution)
True α_2 = -0.5
True α_3 = 0.5

True ρ	Sample size	Par.	d.f. = 5 t-Dist.	d.f. = 5 normal	d.f. = 10 t-Dist.	d.f. = 10 normal
0.0	20	$\alpha(2)$	0.2646	0.2977	0.2730	0.3469
		$\alpha(3)$	0.2637	0.3553	0.2687	0.5656
		ρ	0.3111	0.3160	0.3057	0.2594
	40	$\alpha(2)$	0.1821	0.2454	0.2208	0.2548
		$\alpha(3)$	0.1590	0.2738	0.1648	0.2563
		ρ	0.2974	0.2090	0.2112	0.1937
	100	$\alpha(2)$	0.1082	0.1673	0.1194	0.1311
		$\alpha(3)$	0.1147	0.1888	0.1579	0.1689
		ρ	0.1844	0.1381	0.1671	0.1202
0.2	20	$\alpha(2)$	0.2748	0.2736	0.3073	0.3317
		$\alpha(3)$	0.2859	0.3782	0.2709	0.6103
		ρ	0.3079	0.2842	0.2554	0.2546
	40	$\alpha(2)$	0.1778	0.2430	0.2129	0.2433
		$\alpha(3)$	0.1716	0.2659	0.1801	0.2559
		ρ	0.2837	0.1958	0.2362	0.1898
	100	$\alpha(2)$	0.1182	0.1743	0.1201	0.1218
		$\alpha(3)$	0.1229	0.1694	0.1446	0.1675
		ρ	0.1675	0.1269	0.1691	0.1101
0.5	20	$\alpha(2)$	0.2439	0.2632	0.2443	0.2904
		$\alpha(3)$	0.2686	0.4572	0.3172	0.3851
		ρ	0.2317	0.2122	0.2473	0.1798
	40	$\alpha(2)$	0.1781	0.2200	0.1927	0.2371
		$\alpha(3)$	0.1858	0.2155	0.1894	0.2431
		ρ	0.1953	0.1521	0.1797	0.1508
	100	$\alpha(2)$	0.1125	0.1647	0.1000	0.1123
		$\alpha(3)$	0.1187	0.1409	0.1282	0.1542
		ρ	0.1378	0.1083	0.1050	0.0938
0.7	20	$\alpha(2)$	0.2021	0.2284	0.1739	0.2295
		$\alpha(3)$	0.2332	0.3170	0.2797	0.3095
		ρ	0.1903	0.1502	0.1790	0.1144
	40	$\alpha(2)$	0.1381	0.1857	0.1315	0.1936
		$\alpha(3)$	0.1749	0.1742	0.1978	0.2069
		ρ	0.1468	0.1116	0.1316	0.1034
	100	$\alpha(2)$	0.0904	0.1263	0.0963	0.1090
		$\alpha(3)$	0.1087	0.1186	0.1137	0.1366
		ρ	0.0967	0.0823	0.0918	0.0872

't-dist.' denotes estimates based on the t distribution assumption.
'normal' denotes estimates based on the normality assumption.

Table 3.3

RMS of Estimates

(t distribution vs normal distribution)

True α_2 = -0.3

True α_3 = 0.7

True ρ	Sample size	Par.	d.f. = 5 t-Dist.	d.f. = 5 normal	d.f. = 10 t-Dist.	d.f. = 10 normal
0.0	20	$\alpha(2)$	0.2058	0.2794	0.2682	0.3638
		$\alpha(3)$	0.2680	0.3840	0.2682	0.5846
		ρ	0.3122	0.3046	0.3381	0.2784
	40	$\alpha(2)$	0.1274	0.2201	0.1802	0.2421
		$\alpha(3)$	0.1944	0.2824	0.1838	0.2505
		ρ	0.2456	0.2165	0.2349	0.1953
	100	$\alpha(2)$	0.0781	0.1334	0.1070	0.1255
		$\alpha(3)$	0.1755	0.1936	0.1468	0.1653
		ρ	0.1947	0.1320	0.1923	0.1239
0.2	20	$\alpha(2)$	0.1978	0.2875	0.2992	0.3497
		$\alpha(3)$	0.2678	0.3811	0.2212	0.5778
		ρ	0.2861	0.2792	0.2822	0.2487
	40	$\alpha(2)$	0.1444	0.2235	0.1748	0.2460
		$\alpha(3)$	0.2034	0.2701	0.1949	0.2512
		ρ	0.2482	0.1977	0.2365	0.1863
	100	$\alpha(2)$	0.0829	0.1300	0.1049	0.1196
		$\alpha(3)$	0.1716	0.1921	0.1503	0.1646
		ρ	0.1494	0.1182	0.1567	0.1229
0.5	20	$\alpha(2)$	0.2020	0.2325	0.2542	0.2939
		$\alpha(3)$	0.2663	0.4537	0.3381	0.6498
		ρ	0.2551	0.2143	0.2679	0.1947
	40	$\alpha(2)$	0.1184	0.1785	0.1818	0.2220
		$\alpha(3)$	0.1837	0.2418	0.1690	0.2431
		ρ	0.2028	0.1521	0.2082	0.1599
	100	$\alpha(2)$	0.0776	0.1220	0.0956	0.1085
		$\alpha(3)$	0.1581	0.1879	0.1336	0.1458
		ρ	0.1239	0.1068	0.1212	0.1010
0.7	20	$\alpha(2)$	0.1586	0.2048	0.2191	0.2444
		$\alpha(3)$	0.2215	0.4038	0.2562	0.3368
		ρ	0.2412	0.1503	0.2490	0.1263
	40	$\alpha(2)$	0.1166	0.1604	0.1581	0.1888
		$\alpha(3)$	0.1974	0.2084	0.1873	0.2324
		ρ	0.1625	0.1202	0.1913	0.0931
	100	$\alpha(2)$	0.0605	0.1107	0.0856	0.0918
		$\alpha(3)$	0.1336	0.1541	0.1131	0.1384
		ρ	0.0911	0.0838	0.0915	0.0619

't-dist.' denotes estimates based on the t distribution assumption.
'normal' denotes estimates based on the normality assumption.

Table 4.1
RMS of Estimates
(contaminated normal vs normal distribution)

True α_2 = -1.0

True α_3 = 0.0

True Corr. = 0.0

N	ϵ	Para.	$\sigma = 0.1$		$\sigma = 0.5$	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	0.4868	0.8504	0.7911	0.8536
		$\alpha(3)$	0.2262	0.3174	0.2912	0.3153
		ρ	0.3107	0.3355	0.3214	0.3354
	0.3	$\alpha(2)$	1.1966	0.8480	0.7721	1.5860
		$\alpha(3)$	0.0836	0.3015	0.2398	0.3075
		ρ	0.3754	0.3981	0.3529	0.3873
	0.5	$\alpha(2)$	1.0130	1.6568	0.8694	1.4031
		$\alpha(3)$	0.0449	0.5157	0.1988	0.3004
		ρ	0.3852	0.4287	0.3615	0.4297
40	0.1	$\alpha(2)$	0.2862	0.3091	0.2860	0.3054
		$\alpha(3)$	0.1534	0.2428	0.2199	0.2426
		ρ	0.1969	0.2030	0.1972	0.2021
	0.3	$\alpha(2)$	1.1679	0.5734	0.5450	0.5686
		$\alpha(3)$	0.0676	0.2372	0.1816	0.2391
		ρ	0.2100	0.2499	0.2061	0.2320
	0.5	$\alpha(2)$	0.4921	0.9514	1.0641	0.6332
		$\alpha(3)$	0.0315	0.2317	0.1521	0.2349
		ρ	0.2253	0.3170	0.2157	0.2696
100	0.1	$\alpha(2)$	0.1380	0.1687	0.1323	0.1530
		$\alpha(3)$	0.0551	0.1325	0.1197	0.1323
		ρ	0.0965	0.1048	0.0968	0.1011
	0.3	$\alpha(2)$	0.1755	0.3059	0.1557	0.2630
		$\alpha(3)$	0.0238	0.1328	0.0997	0.1324
		ρ	0.0951	0.1302	0.0981	0.1146
	0.5	$\alpha(2)$	0.1971	0.4419	0.1434	0.3638
		$\alpha(3)$	0.0461	0.1326	0.0862	0.1324
		ρ	0.0909	0.1652	0.0967	0.1262

'Cont.-N' denotes the estimates based on the contaminated normal assumption.
'Normal' denotes the estimates based on the normality assumption.

<u>Table 4.2</u>
<u>RMS of Estimates</u>
<u>(contaminated normal vs normal distribution)</u>

True α_2 = -1.0
True α_3 = 0.0
True Corr. = 0.2

N	ϵ	Para.	σ = 0.1		σ = 0.5	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	0.8783	0.8495	0.8758	0.8508
		$\alpha(3)$	0.2392	0.3244	0.2996	0.3229
		ρ	0.2808	0.3018	0.2917	0.2990
	0.3	$\alpha(2)$	1.2174	0.7749	0.8137	1.0633
		$\alpha(3)$	0.0996	0.3064	0.2482	0.3141
		ρ	0.3222	0.3863	0.3225	0.3497
	0.5	$\alpha(2)$	1.1853	1.4648	0.8229	2.1362
		$\alpha(3)$	0.0457	0.3002	0.2109	0.3057
		ρ	0.3281	0.4058	0.3267	0.3796
40	0.1	$\alpha(2)$	0.2855	0.3086	0.2856	0.3056
		$\alpha(3)$	0.1582	0.2560	0.2335	0.2572
		ρ	0.1938	0.2001	0.1954	0.1987
	0.3	$\alpha(2)$	0.3115	0.4021	0.3039	0.3893
		$\alpha(3)$	0.1062	0.2500	0.1937	0.2538
		ρ	0.2135	0.2514	0.2081	0.2277
	0.5	$\alpha(2)$	0.3016	0.4847	0.2790	0.4658
		$\alpha(3)$	0.0341	0.2459	0.1633	0.2515
		ρ	0.2181	0.3218	0.2164	0.2616
100	0.1	$\alpha(2)$	0.1509	0.1790	0.1495	0.1692
		$\alpha(3)$	0.0572	0.1268	0.1151	0.1267
		ρ	0.0913	0.0970	0.0937	0.0954
	0.3	$\alpha(2)$	0.1859	0.3020	0.1685	0.2687
		$\alpha(3)$	0.0257	0.1253	0.0960	0.1257
		ρ	0.0978	0.1275	0.1006	0.1113
	0.5	$\alpha(2)$	0.1916	0.4130	0.1483	0.3576
		$\alpha(3)$	0.0693	0.1248	0.0825	0.1250
		ρ	0.0971	0.1722	0.1016	0.1314

'Cont.-N' denotes the estimates based on the contaminated normal assumption.
'Normal' denotes the estimates based on the normality assumption.

Table 4.3

RMS of Estimates

(contaminated normal vs normal distribution)

True α_2 = -1.0

True α_3 = 0.0

True Corr. = 0.5

N	ϵ	Para.	σ = 0.1		σ = 0.5	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	0.8445	0.9431	0.9085	0.9009
		$\alpha(3)$	0.2279	0.2957	0.2762	0.2971
		ρ	0.1785	0.1922	0.1834	0.1873
	0.3	$\alpha(2)$	1.0437	1.3703	1.4003	1.0366
		$\alpha(3)$	0.0662	0.2689	0.2273	0.2808
		ρ	0.1993	0.2290	0.2014	0.2101
	0.5	$\alpha(2)$	1.4617	0.9916	1.1620	0.9856
		$\alpha(3)$	0.0470	0.2550	0.1954	0.2720
		ρ	0.2127	0.2730	0.2147	0.2455
40	0.1	$\alpha(2)$	0.2898	0.3088	0.3001	0.3122
		$\alpha(3)$	0.1273	0.1988	0.1831	0.1994
		ρ	0.1449	0.1445	0.1410	0.1328
	0.3	$\alpha(2)$	1.0729	0.5808	0.4078	0.5375
		$\alpha(3)$	0.0545	0.1937	0.1533	0.1952
		ρ	0.1576	0.1621	0.1492	0.1462
	0.5	$\alpha(2)$	0.4575	0.5066	0.5692	0.5538
		$\alpha(3)$	0.0296	0.1855	0.1310	0.1890
		ρ	0.1667	0.2464	0.1659	0.1882
100	0.1	$\alpha(2)$	0.1320	0.1518	0.1277	0.1454
		$\alpha(3)$	0.0477	0.1047	0.0953	0.1035
		ρ	0.0772	0.0797	0.0763	0.0778
	0.3	$\alpha(2)$	0.1495	0.2299	0.1346	0.2149
		$\alpha(3)$	0.0225	0.0994	0.0789	0.0992
		ρ	0.0836	0.1042	0.0819	0.0969
	0.5	$\alpha(2)$	0.1574	0.2961	0.1179	0.2720
		$\alpha(3)$	0.0537	0.0965	0.0687	0.0961
		ρ	0.0895	0.1476	0.0846	0.1281

'Cont.-N' denotes the estimates based on the contaminated normal assumption.
'Normal' denotes the estimates based on the normality assumption.

Table 4.4
RMS of Estimates
(contaminated normal vs normal distribution)

True α_2 = -1.0

True α_3 = 0.0

True Corr. = 0.7

N	ϵ	Para.	σ = 0.1		σ = 0.5	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	1.1364	0.8128	0.8635	0.8185
		$\alpha(3)$	0.1864	0.2536	0.2390	0.2550
		ρ	0.1219	0.1249	0.1157	0.1167
	0.3	$\alpha(2)$	1.4245	0.7298	0.8054	0.8116
		$\alpha(3)$	0.0648	0.2199	0.2004	0.2345
		ρ	0.1391	0.1386	0.1367	0.1261
	0.5	$\alpha(2)$	1.1189	3.0079	0.8395	0.8853
		$\alpha(3)$	0.0418	0.1951	0.1718	0.2127
		ρ	0.1450	0.1324	0.1497	0.1271
40	0.1	$\alpha(2)$	0.2652	0.2616	0.2599	0.2669
		$\alpha(3)$	0.1124	0.1769	0.1656	0.1780
		ρ	0.1018	0.1038	0.1012	0.1003
	0.3	$\alpha(2)$	0.2691	0.2905	0.2608	0.3011
		$\alpha(3)$	0.0510	0.1661	0.1398	0.1731
		ρ	0.1016	0.1054	0.1040	0.0993
	0.5	$\alpha(2)$	0.2610	0.3089	0.2308	0.2974
		$\alpha(3)$	0.0269	0.1513	0.1200	0.1610
		ρ	0.1115	0.1499	0.1152	0.1155
100	0.1	$\alpha(2)$	0.1298	0.1294	0.1186	0.1277
		$\alpha(3)$	0.0466	0.0852	0.0715	0.0833
		ρ	0.0561	0.0591	0.0571	0.0574
	0.3	$\alpha(2)$	0.1496	0.1725	0.1260	0.1671
		$\alpha(3)$	0.0219	0.0808	0.0647	0.0790
		ρ	0.0559	0.0782	0.0619	0.0736
	0.5	$\alpha(2)$	0.1572	0.1921	0.1138	0.1948
		$\alpha(3)$	0.0628	0.0725	0.0560	0.0728
		ρ	0.0652	0.1048	0.0627	0.0951

'Cont.-N' denotes the estimates based on the contaminated normal assumption.
'Normal' denotes the estimates based on the normality assumption.

Table 4.5
RMS of Estimates
(contaminated normal vs normal distribution)

True α_2 = -0.5

True α_3 = 0.5

True Corr. = 0.0

N	ϵ	Para.	$\sigma = 0.1$		$\sigma = 0.5$	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	0.3098	0.3349	0.2729	0.2924
		$\alpha(3)$	0.3531	0.3859	0.3360	0.3659
		ρ	0.2601	0.2800	0.2676	0.2743
	0.3	$\alpha(2)$	0.3402	0.4270	0.2498	0.3220
		$\alpha(3)$	0.3952	0.5525	0.3410	0.4475
		ρ	0.3150	0.3629	0.2818	0.3035
	0.5	$\alpha(2)$	0.3153	0.5105	0.2264	0.3595
		$\alpha(3)$	1.0652	1.0645	0.3082	0.4961
		ρ	0.3995	0.4615	0.3044	0.3602
40	0.1	$\alpha(2)$	0.2086	0.2369	0.1803	0.2048
		$\alpha(3)$	0.2452	0.2637	0.2238	0.2492
		ρ	0.1970	0.2069	0.1995	0.2045
	0.3	$\alpha(2)$	0.2401	0.3851	0.1764	0.2692
		$\alpha(3)$	0.2878	0.4347	0.2078	0.3067
		ρ	0.2199	0.2569	0.2076	0.2276
	0.5	$\alpha(2)$	0.2725	0.5497	0.1590	0.3295
		$\alpha(3)$	0.3595	0.6693	0.1846	0.3752
		ρ	0.2950	0.3696	0.2252	0.2731
100	0.1	$\alpha(2)$	0.1298	0.1641	0.1098	0.1306
		$\alpha(3)$	0.1201	0.1534	0.1083	0.1280
		ρ	0.1053	0.1115	0.1056	0.1087
	0.3	$\alpha(2)$	0.1411	0.3248	0.0927	0.1848
		$\alpha(3)$	0.1507	0.3196	0.1091	0.1985
		ρ	0.1140	0.1355	0.1058	0.1168
	0.5	$\alpha(2)$	0.1400	0.5159	0.0802	0.2633
		$\alpha(3)$	0.1874	0.5509	0.1079	0.2897
		ρ	0.1494	0.2064	0.1110	0.1353

'Cont.-N' denotes the estimates based on the contaminated normal assumption
'Normal' denotes the estimates based on the normality assumption.

Table 4.6
RMS of Estimates
(contaminated normal vs normal distribution)

True α_2 = -0.5

True α_3 = 0.5

True Corr. = 0.2

N	ϵ	Para.	σ = 0.1		σ = 0.5	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	0.2796	0.3116	0.2678	0.2851
		$\alpha(3)$	0.3103	0.3525	0.2995	0.3300
		ρ	0.2415	0.2614	0.2492	0.2553
	0.3	$\alpha(2)$	0.2979	0.3967	0.2255	0.2903
		$\alpha(3)$	0.5734	0.7986	0.3175	0.4212
		ρ	0.2839	0.3318	0.2575	0.2791
	0.5	$\alpha(2)$	0.2742	0.4862	0.2114	0.3271
		$\alpha(3)$	1.0369	0.8730	0.3086	0.4960
		ρ	0.3662	0.4439	0.2801	0.3312
40	0.1	$\alpha(2)$	0.2075	0.2203	0.1713	0.1909
		$\alpha(3)$	0.2437	0.2585	0.2212	0.2494
		ρ	0.1831	0.1928	0.1796	0.1834
	0.3	$\alpha(2)$	0.2265	0.3567	0.1655	0.2437
		$\alpha(3)$	0.2621	0.4078	0.2024	0.3043
		ρ	0.2173	0.2568	0.1975	0.2168
	0.5	$\alpha(2)$	0.2578	0.5119	0.1481	0.2942
		$\alpha(3)$	0.3134	0.6179	0.1830	0.3690
		ρ	0.2655	0.3497	0.2101	0.2568
100	0.1	$\alpha(2)$	0.1276	0.1628	0.1052	0.1252
		$\alpha(3)$	0.1290	0.1533	0.1145	0.1316
		ρ	0.1005	0.1046	0.0981	0.1006
	0.3	$\alpha(2)$	0.1492	0.3189	0.0936	0.1781
		$\alpha(3)$	0.1490	0.3025	0.1126	0.1932
		ρ	0.1120	0.1329	0.0985	0.1093
	0.5	$\alpha(2)$	0.1605	0.4872	0.0818	0.2441
		$\alpha(3)$	0.1302	0.5198	0.1099	0.2794
		ρ	0.1412	0.2003	0.0987	0.1287

'Cont.-N' denotes the estimates based on the contaminated normal assumption
'Normal' denotes the estimates based on the normality assumption.

Table 4.7
RMS of Estimates
(contaminated normal vs normal distribution)

True α_2 = -0.5

True α_3 = 0.5

True Corr. = 0.5

N	ϵ	Para.	σ = 0.1		σ = 0.5	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	0.2863	0.2999	0.2635	0.2801
		$\alpha(3)$	0.2728	0.3049	0.2582	0.2821
		ρ	0.1873	0.1893	0.1952	0.1978
	0.3	$\alpha(2)$	0.2878	0.3423	0.2321	0.2850
		$\alpha(3)$	0.3677	0.4762	0.2654	0.3378
		ρ	0.2342	0.2507	0.2098	0.2089
	0.5	$\alpha(2)$	0.2819	0.4276	0.2252	0.3047
		$\alpha(3)$	0.6459	0.6943	0.2728	0.4189
		ρ	0.3379	0.3292	0.2210	0.2473
40	0.1	$\alpha(2)$	0.1998	0.2025	0.1790	0.1923
		$\alpha(3)$	0.2057	0.2348	0.2012	0.2240
		ρ	0.1401	0.1471	0.1378	0.1394
	0.3	$\alpha(2)$	0.2216	0.3086	0.1726	0.2273
		$\alpha(3)$	0.2301	0.3557	0.1771	0.2596
		ρ	0.1740	0.1988	0.1543	0.1649
	0.5	$\alpha(2)$	0.2520	0.4117	0.1508	0.2413
		$\alpha(3)$	0.2985	0.5228	0.1757	0.3177
		ρ	0.2106	0.2549	0.1531	0.1879
100	0.1	$\alpha(2)$	0.1121	0.1360	0.0937	0.1055
		$\alpha(3)$	0.1229	0.1388	0.1061	0.1194
		ρ	0.0853	0.0897	0.0829	0.0845
	0.3	$\alpha(2)$	0.1415	0.2627	0.0924	0.1524
		$\alpha(3)$	0.1334	0.2505	0.0925	0.1528
		ρ	0.0997	0.1307	0.0873	0.1038
	0.5	$\alpha(2)$	0.1528	0.3753	0.0752	0.1871
		$\alpha(3)$	0.1504	0.4039	0.0834	0.2098
		ρ	0.1199	0.1827	0.0879	0.1316

'Cont.-N' denotes the estimates based on the contaminated normal assumption
'Normal' denotes the estimates based on the normality assumption.

Table 4.8

RMS of Estimates

(contaminated normal vs normal distribution)

True α_2 = -0.5

True α_3 = 0.5

True Corr. = 0.7

N	ϵ	Para.	σ = 0.1		σ = 0.5	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	0.2380	0.2330	0.2201	0.42255
		$\alpha(3)$	0.2611	0.2733	0.2410	0.2589
		ρ	0.1219	0.1204	0.1278	0.1245
	0.3	$\alpha(2)$	0.2309	0.2370	0.1947	0.2168
		$\alpha(3)$	0.4786	0.3578	0.2537	0.3022
		ρ	0.1611	0.1546	0.1468	0.1418
	0.5	$\alpha(2)$	0.2655	0.2990	0.2083	0.2353
		$\alpha(3)$	0.5370	0.4397	0.2304	0.3166
		ρ	0.1868	0.1942	0.1508	0.1471
40	0.1	$\alpha(2)$	0.1858	0.1714	0.1723	0.1775
		$\alpha(3)$	0.1819	0.2016	0.1768	0.1942
		ρ	0.0889	0.0950	0.0857	0.0886
	0.3	$\alpha(2)$	0.2084	0.2408	0.1627	0.1877
		$\alpha(3)$	0.2054	0.2739	0.1538	0.2072
		ρ	0.1090	0.1279	0.0936	0.1054
	0.5	$\alpha(2)$	0.2488	0.3172	0.1522	0.1991
		$\alpha(3)$	0.2448	0.3874	0.1584	0.2421
		ρ	0.1334	0.1527	0.1008	0.1248
100	0.1	$\alpha(2)$	0.1025	0.1108	0.0905	0.0943
		$\alpha(3)$	0.1035	0.1159	0.0917	0.1010
		ρ	0.0498	0.0572	0.0499	0.0526
	0.3	$\alpha(2)$	0.1322	0.2017	0.0850	0.1205
		$\alpha(3)$	0.1215	0.1906	0.0807	0.1179
		ρ	0.0611	0.0946	0.0536	0.0732
	0.5	$\alpha(2)$	0.1512	0.2659	0.0741	0.1364
		$\alpha(3)$	0.1258	0.2783	0.0729	0.1469
		ρ	0.0764	0.1329	0.0588	0.1020

'Cont.-N' denotes the estimates based on the contaminated normal assumption

'Normal' denotes the estimates based on the normality assumption.

Table 4.9
RMS of Estimates
(contaminated normal vs normal distribution)

True α_2 = -0.3

True α_3 = 0.7

True Corr. = 0.0

N	ϵ	Para.	$\sigma = 0.1$		$\sigma = 0.5$	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	0.2784	0.3228	0.3065	0.3243
		$\alpha(3)$	0.3439	0.3762	0.3344	0.3610
		ρ	0.3057	0.3251	0.3129	0.3214
	0.3	$\alpha(2)$	0.2601	0.4111	0.2593	0.3134
		$\alpha(3)$	0.8821	0.8912	0.3699	0.4699
		ρ	0.3426	0.3797	0.3255	0.3496
	0.5	$\alpha(2)$	0.2469	0.5578	0.2286	0.3205
		$\alpha(3)$	1.7282	1.5748	0.5482	0.6447
		ρ	0.3861	0.4292	0.3397	0.3942
40	0.1	$\alpha(2)$	0.1843	0.2308	0.2038	0.2207
		$\alpha(3)$	0.2683	0.2938	0.2570	0.2819
		ρ	0.2005	0.2128	0.1964	0.2019
	0.3	$\alpha(2)$	0.1808	0.3735	0.1658	0.2201
		$\alpha(3)$	0.2771	0.4113	0.2426	0.3457
		ρ	0.2280	0.2728	0.2073	0.2315
	0.5	$\alpha(2)$	0.1819	0.5660	0.1485	0.2433
		$\alpha(3)$	0.3461	0.6260	0.2648	0.4722
		ρ	0.2747	0.3531	0.2086	0.2575
100	0.1	$\alpha(2)$	0.1187	0.1568	0.1077	0.1185
		$\alpha(3)$	0.1356	0.1556	0.1224	0.1401
		ρ	0.1029	0.1075	0.0985	0.1015
	0.3	$\alpha(2)$	0.1209	0.3445	0.0920	0.1438
		$\alpha(3)$	0.1392	0.2777	0.1119	0.2087
		ρ	0.1082	0.1268	0.0984	0.1084
	0.5	$\alpha(2)$	0.1073	0.5531	0.0790	0.1792
		$\alpha(3)$	0.1636	0.4879	0.1143	0.3389
		ρ	0.1276	0.1917	0.0993	0.1229

'Cont.-N' denotes estimates based on the contaminated normal assumption.
'Normal' denotes estimates based on the normality assumption.

Table 4.10

RMS of Estimates

(contaminated normal vs normal distribution)

True α_2 = -0.3

True α_3 = 0.7

True Corr. = 0.2

N	ϵ	Para.	σ = 0.1		σ = 0.5	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	0.2781	0.3269	0.3048	0.3235
		$\alpha(3)$	0.3816	0.4062	0.3705	0.3960
		ρ	0.2766	0.2960	0.2864	0.2934
	0.3	$\alpha(2)$	0.2788	0.4340	0.2664	0.3263
		$\alpha(3)$	1.0751	0.7734	0.3714	0.4687
		ρ	0.3121	0.3460	0.2877	0.3123
	0.5	$\alpha(2)$	0.2521	0.5688	0.2312	0.3282
		$\alpha(3)$	1.0227	1.1392	0.4993	0.6205
		ρ	0.3873	0.4020	0.2978	0.3571
40	0.1	$\alpha(2)$	0.1941	0.2450	0.2123	0.2300
		$\alpha(3)$	0.2671	0.2931	0.2606	0.2853
		ρ	0.1858	0.1960	0.1800	0.1843
	0.3	$\alpha(2)$	0.2018	0.3854	0.1770	0.2359
		$\alpha(3)$	0.2674	0.3997	0.2388	0.3413
		ρ	0.2137	0.2548	0.1883	0.2083
	0.5	$\alpha(2)$	0.2009	0.5489	0.1524	0.2489
		$\alpha(3)$	0.2971	0.5723	0.2572	0.4601
		ρ	0.2747	0.3611	0.1878	0.2385
100	0.1	$\alpha(2)$	0.1254	0.1619	0.1099	0.1199
		$\alpha(3)$	0.1391	0.1593	0.1265	0.1452
		ρ	0.0970	0.0997	0.0921	0.0943
	0.3	$\alpha(2)$	0.1359	0.3432	0.0965	0.1428
		$\alpha(3)$	0.1434	0.2153	0.1179	0.2128
		ρ	0.0988	0.1222	0.0906	0.0996
	0.5	$\alpha(2)$	0.1091	0.5314	0.0856	0.1762
		$\alpha(3)$	0.1580	0.4647	0.1189	0.3322
		ρ	0.1213	0.1990	0.0952	0.1247

'Cont.-N' denotes estimates based on the contaminated normal assumption.
'Normal' denotes estimates based on the normality assumption.

Table 4.11
RMS of Estimates
(contaminated normal vs normal distribution)

True α_2 = -0.3

True α_3 = 0.7

True Corr. = 0.5

N	ϵ	Para.	σ = 0.1		σ = 0.5	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	0.2634	0.3005	0.2740	0.2888
		$\alpha(3)$	0.3666	0.3744	0.3507	0.3695
		ρ	0.1894	0.1943	0.2066	0.2061
	0.3	$\alpha(2)$	0.2178	0.3354	0.2341	0.2141
		$\alpha(3)$	0.8399	0.4455	0.3556	0.4281
		ρ	0.2482	0.2556	0.2366	0.2338
	0.5	$\alpha(2)$	0.2247	0.4587	0.1902	0.2588
		$\alpha(3)$	2.0017	1.0009	0.4531	0.6941
		ρ	0.3317	0.3124	0.2708	0.3025
40	0.1	$\alpha(2)$	0.1964	0.2284	0.2082	0.2240
		$\alpha(3)$	0.2527	0.2673	0.2410	0.2614
		ρ	0.1514	0.1586	0.1448	0.1482
	0.3	$\alpha(2)$	0.2177	0.3476	0.1840	0.2292
		$\alpha(3)$	0.2875	0.3632	0.2258	0.3119
		ρ	0.1811	0.2056	0.1533	0.1684
	0.5	$\alpha(2)$	0.2279	0.4779	0.1567	0.2283
		$\alpha(3)$	0.5583	0.6400	0.2348	0.3917
		ρ	0.2444	0.3002	0.1704	0.2030
100	0.1	$\alpha(2)$	0.1229	0.1469	0.1083	0.1162
		$\alpha(3)$	0.1301	0.1422	0.1191	0.1318
		ρ	0.0734	0.0818	0.0710	0.0748
	0.3	$\alpha(2)$	0.1368	0.3013	0.1000	0.1340
		$\alpha(3)$	0.1382	0.2191	0.1152	0.1793
		ρ	0.0883	0.1242	0.0742	0.0943
	0.5	$\alpha(2)$	0.1208	0.4424	0.0871	0.1531
		$\alpha(3)$	0.1549	0.3365	0.1193	0.2546
		ρ	0.1145	0.1800	0.0836	0.1266

'Cont.-N' denotes estimates based on the contaminated normal assumption.
'Normal' denotes estimates based on the normality assumption.

Table 4.12

RMS of Estimates

(contaminated normal vs normal distribution)

True α_2 = -0.3

True α_3 = 0.7

True Corr. = 0.7

N	ϵ	Para.	σ = 0.1		σ = 0.5	
			Cont.-N.	Normal	Cont.-N.	Normal
20	0.1	$\alpha(2)$	0.2170	0.2337	0.2226	0.2331
		$\alpha(3)$	0.4115	0.4244	0.4008	0.4426
		ρ	0.1192	0.1224	0.1202	0.1171
	0.3	$\alpha(2)$	0.1976	0.2917	0.2032	0.2309
		$\alpha(3)$	1.6634	0.4960	0.3893	0.4429
		ρ	0.1438	0.1556	0.1530	0.1505
	0.5	$\alpha(2)$	0.2015	0.3526	0.1657	0.2071
		$\alpha(3)$	0.5983	0.5647	0.3547	0.3464
		ρ	0.1531	0.1705	0.1695	0.1529
40	0.1	$\alpha(2)$	0.1603	0.1792	0.1617	0.1724
		$\alpha(3)$	0.2460	0.2392	0.2247	0.2367
		ρ	0.0936	0.0997	0.0925	0.0939
	0.3	$\alpha(2)$	0.1868	0.2873	0.1483	0.1788
		$\alpha(3)$	0.2344	0.2704	0.2071	0.2511
		ρ	0.1093	0.1253	0.0999	0.1094
	0.5	$\alpha(2)$	0.1965	0.3925	0.1244	0.1743
		$\alpha(3)$	0.2540	0.5216	0.4301	0.4407
		ρ	0.1324	0.1578	0.1083	0.1275
100	0.1	$\alpha(2)$	0.1000	0.1232	0.0906	0.0951
		$\alpha(3)$	0.1248	0.1273	0.1149	0.1223
		ρ	0.0532	0.0608	0.0530	0.0561
	0.3	$\alpha(2)$	0.1165	0.2453	0.0876	0.1088
		$\alpha(3)$	0.1348	0.1640	0.1123	0.1463
		ρ	0.0610	0.0965	0.0551	0.0746
	0.5	$\alpha(2)$	0.1054	0.3498	0.0730	0.1159
		$\alpha(3)$	0.1538	0.2300	0.1159	0.1907
		ρ	0.0797	0.1294	0.0604	0.0969

'Cont.-N' denotes estimates based on the contaminated normal assumption.
'Normal' denotes estimates based on the normality assumption.

Appendix A

The left hand side of (2.17) is given as

$$t_2(X,Y;V,n) = \frac{\Gamma[(n+2)/2]}{\Gamma(n/2)\cdot(n\pi)} |V|^{-1/2} [1 + n^{-1}(X\ Y)\ V^{-1}\ (X\ Y)']^{-\frac{n+2}{2}} .$$

This joint density function can be decomposed into products of two density functions as follows.

$$t_2(X,Y;V,n) = (2\pi)^{-1}(\frac{n}{n-2})(1-\rho^2)^{-1/2}[1 + \frac{X^2 - 2\rho X Y + Y^2}{(n-2)\ (1-\rho^2)}]^{-\frac{n+2}{2}}$$

$$= (2\pi)^{-1}(\frac{n}{n-2})(1-\rho^2)^{-1/2}[1 + \frac{X^2\ (1-\rho^2) + (Y - \rho X)^2}{(n-2)\ (1-\rho^2)}]^{-\frac{n+2}{2}}$$

$$= (2\pi)^{-1}(\frac{n}{n-2})(1-\rho^2)^{-1/2}[1 + \frac{X^2}{n-2} + \frac{(Y - \rho X)^2}{(n-2)\ (1-\rho^2)}]^{-\frac{n+2}{2}} .$$

If we take $\nu_X = \frac{n-2}{n+1}(1-\rho^2)(1+\frac{X^2}{n-2})$, then $t_2(X,Y;V,n)$ become

$$(2\pi)^{-1}(\frac{n}{n-2})(1-\rho^2)^{-1/2}[1 + \frac{X^2}{n-2} + \frac{(Y - \rho X)^2}{(n+1)\ \nu_X}(1+\frac{X^2}{n-2})]^{-\frac{n+2}{2}}$$

$$= (2\pi)^{-1}(\frac{n}{n-2})[(1-\rho^2)(1 + \frac{X^2}{n-2})]^{-1/2}$$

$$\times [1 + \frac{(Y - \rho X)^2}{(n+1)\ \nu_X}]^{-\frac{n+2}{2}}(1 + \frac{X^2}{n-2})^{-\frac{n+1}{2}}$$

$$= (2\pi)^{-1}n[(n+1)(n-2)]^{-1/2}\nu_X^{-1/2}[1 + \frac{X^2}{n-2}]^{-\frac{n+1}{2}}[1 + \frac{(Y-\rho X)^2}{(n+1)\cdot\nu_X}]^{-\frac{n+2}{2}}$$

$$= \frac{\Gamma[(n+2)/2]}{\Gamma(n/2)(n\pi)^{1/2}[(n+1)\pi]^{1/2}}\left(\frac{n}{n-2}\right)^{1/2} \nu_x^{-1/2}\left[1+\frac{X^2}{n-2}\right]^{-\frac{n+1}{2}}\left[1+\frac{(Y-\rho X)^2}{(n+1)\nu_x}\right]^{-\frac{n+2}{2}}$$

$$= \frac{\Gamma[(n+1)/2]}{\Gamma(n/2)\ (n\pi)^{1/2}}\left(\frac{n}{n-2}\right)^{1/2}\left[1+\frac{1}{n}\frac{X^2}{\left(\frac{n-2}{n}\right)}\right]^{-\frac{n+1}{2}}$$

$$\times \frac{\Gamma[(n+2)/2]}{\Gamma[(n+1)/2]\ [(n+1)\pi)]^{1/2}}\ \nu_x^{-1/2}\left[1+\frac{1}{n+1}\frac{(Y-\rho X)^2}{\nu_x}\right]^{-\frac{n+2}{2}}$$

$$= t_1(X;0,\frac{n-2}{n},n) \times t_1(Y;\rho X,\nu_x,n+1).$$

That is, the joint density function of $(X,Y)'$, $t_2(X,Y;V,n)$, has been decomposed into a product of the marginal density function of X and the conditional density function of Y given X=x.

Appendix B

Since $\frac{\partial \text{ Log } L(\underline{\theta})}{\partial \rho}$ involves the partial derivative of

$$\frac{\partial}{\partial\rho}\int_{-\infty}^{\alpha_{z_i}} t_1(Y;\rho x_i, \frac{n-2}{n+1}(1+\frac{x_i^2}{n-2})(1-\rho^2), n+1)dY, \qquad \text{(B1)}$$

we give the expression of its partial derivative as belows first.

Let $\mu_x=\rho x_i$, $\nu_x = \frac{n-2}{n+1}(1+\frac{x_i^2}{n-2})(1-\rho^2)$, $m = n+1$ and $\alpha = \alpha_{z_i}$, then the expression of (B1) become

$$\frac{\partial}{\partial\rho}\int_{-\infty}^{\alpha} t_1(Y;\mu_x,\nu_x,m)dY$$

$$= \frac{\partial}{\partial\rho}\int_{-\infty}^{\alpha^*} t_1(\omega;0,1,m)d\omega, \qquad \text{(B2)}$$

where $\alpha^* = (\alpha - \mu_x)\nu_x^{-1/2}$ and $\omega = (Y-\mu_x)\nu_x^{-1/2}$.

The integral of (B2) depends only on α^*, which is, in turn, a function of ρ. Thus, the expression of (B2) can be obtained by using chain rule as follows.

$$\frac{\partial}{\partial\rho}\int_{-\infty}^{\alpha} t_1(Y;\mu_x,\nu_x,m)dY = \frac{\partial}{\partial\alpha^*}\int_{-\infty}^{\alpha^*} t_1(\omega;0,1,m)d\omega \cdot \frac{\partial\alpha^*}{\partial\rho}$$

$$= t_1(\alpha^*;0,1,m) \cdot \frac{\partial}{\partial\rho}\{(\alpha-\rho\, x_i)\cdot[\frac{n-2}{n+1}(1+\frac{x_i^2}{n-2})(1-\rho^2)]^{-1/2}\}$$

$$= t_1(\alpha^*;0,1,m)\quad [\frac{n-2}{n+1}(1+\frac{x_i^2}{n-2})]^{-1/2} \quad \frac{\rho\cdot\alpha - x_i}{(1-\rho^2)^{3/2}}$$

$$= \frac{t_1(\alpha^*;0,1,m)}{\sqrt{\nu_x}}\ \frac{\rho\cdot\alpha - x_i}{(1-\rho^2)}$$

As a result, the expression of $\frac{\partial \text{ Log } L(\underline{\theta})}{\partial \rho}$ can be obtained and given as

$$\frac{\partial \text{ Log } L(\underline{\theta})}{\partial \rho} = \sum_{i=1}^{N} \frac{1}{T(\alpha_{z_i+1} | X=x_i) - T(\alpha_{z_i} | X=x_i)}$$

$$\times \left\{ \frac{t_1(\alpha^*_{z_i+1}; 0, 1, n+1)}{\sqrt{\nu_x}} \cdot \frac{\rho \cdot \alpha_{z_i+1} - x_i}{(1-\rho^2)} - \frac{t_1(\alpha^*_{z_i}; 0, 1, n+1)}{\sqrt{\nu_x}} \cdot \frac{\rho \cdot \alpha_{z_i} - x_i}{(1-\rho^2)} \right\},$$

Appendix C

We are going to show that $\underline{T} = \sqrt{n}\, U^{-1/2}\, V^{1/2}\, \underline{X}$ has a bivariate elliptical t distribution with density function

$$t_2(\underline{T};V,n) = (2\pi)^{-1}\, |V|^{-1/2}\, [\,1 + n^{-1}\, \underline{T}'V^{-1}\underline{T}\,]^{-\frac{n+2}{2}}. \qquad (C1)$$

First, we show that $\underline{\omega} = \sqrt{n}\, U^{-1/2}\, \underline{X}$ is distributed as a bivariate t distribution with density function

$$t_2(\underline{\omega};I_2,n) = (2\pi)^{-1}\, [1 + n^{-1}\underline{\omega}'\underline{\omega}]^{-\frac{n+2}{2}}, \qquad (C2)$$

where $\underline{X} = (X_1, X_2)'$ which is distributed as $N_2(\underline{0}, I_2)$

and $\underline{\omega} = (\omega_1, \omega_2)' = (\sqrt{n}\, U^{-1/2}X_1,\ \sqrt{n}\, U^{-1/2}X_2)'$.

$\underline{X}$ and U are stochastically independent, thus, the joint density function of $(X_1, X_2, U)'$ is

$$g(X_1,X_2,U) = (2\pi)^{-1}\exp\left(-\frac{X_1^2 + X_2^2}{2}\right) \frac{1}{\Gamma(n/2)\, 2^{n/2}}\, U^{n/2-1}\, \exp(-U/2).$$

In addition, since $\omega_1 = \sqrt{n}\, U^{-1/2}X_1$ and $\omega_2 = \sqrt{n}\, U^{-1/2}X_2$, the joint density function of the new variates ω_1, ω_2 and U is obtained as

$$h(\omega_1,\omega_2,U) = (2\pi)^{-1}\, \exp\left[-\frac{U(\omega_1^2 + \omega_2^2)}{2n}\right] \times \frac{1}{\Gamma(n/2)\, 2^{n/2}}\, U^{n/2-1}\, \exp\left(-\frac{U}{2}\right) \cdot \left(\frac{U}{n}\right) \qquad (C3)$$

$$= (2\pi n)^{-1} \frac{1}{\Gamma(n/2)\, 2^{n/2}} U^{n/2} \exp[-\frac{U}{2}(1 + \frac{\omega_1^2 - \omega_2^2}{n})]$$

with $-\infty < \omega_1 < \infty$, $-\infty < \omega_2 < \infty$, and $0 \le U < \infty$.

The last product term of the expression of (C3), $\frac{U}{n}$, is Jacobian. The joint marginal distribution of $(\omega_1, \omega_2)'$ is

$$f(\omega_1, \omega_2) = \int_0^\infty h(\omega_1, \omega_2, U)\, dU$$

$$= (2\pi n)^{-1} \frac{1}{\Gamma(n/2)\, 2^{n/2}} \int_0^\infty U^{n/2} \exp[-\frac{U}{2}(1 + \frac{\omega_1^2 - \omega_2^2}{n})]dU$$

$$= (2\pi n)^{-1} \frac{1}{\Gamma(n/2)\, 2^{n/2}} [1 + \frac{\omega_1^2 + \omega_2^2}{n}]^{-(\frac{n}{2} + 1)}$$

$$\times \int_0^\infty [U(1 + \frac{\omega_1^2 + \omega_2^2}{n})]^{n/2} \exp(-\frac{U}{2}(1 + \frac{\omega_1^2 + \omega_2^2}{n})] \cdot$$

$$d[U(1 + \frac{\omega_1^2 - \omega_2^2}{n})]. \qquad \text{(C4)}$$

The integral of the expression of (C4) is in the form $\int_0^\infty Y^{n/2} \exp(-\frac{Y}{2})\, dY$, which is, in fact, the Gamma function integral,

$$\text{so } f(\omega_1, \omega_2) = (2\pi n)^{-1} \frac{1}{\Gamma(n/2)\, 2^{n/2}} [1 + \frac{\omega_1^2 + \omega_2^2}{n}]^{-\frac{n+2}{2}} \Gamma(\frac{n-2}{2})\, 2^{\frac{n-2}{2}}$$

$$= (2\pi)^{-1} [1 + \frac{\omega_1^2 + \omega_2^2}{n}]^{-\frac{n+2}{2}}.$$

That is, $(\omega_1, \omega_2)'$ is distributed as a bivariate t distribution with zero mean vector and covariance parameter matrix I_2.

If we take $\underline{T} = V^{-1/2}\underline{\omega}$, then it is well-known that $\underline{T}$ is also distributed as a bivariate t distribution with zero mean vector and covariance parameter matrix V.

Since $\underline{T} = V^{1/2}\underline{\omega}$ can be rewritten as $\underline{\omega} = V^{-1/2}\underline{T}$, the Jacobian of the transformation is $|V|^{-1/2}$, the density function of new variate $\underline{T}$ is

$$(2\pi)^{-1}|V|^{-1/2}\,[1 + n^{-1}\underline{T}'V^{-1}\underline{T}]^{-\frac{n+2}{2}}.$$

Hence, $\underline{T}$ is distributed as the desired bivariate elliptical t distribution.

Appendix D

For a correlation matrix $A = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$, we want to find a 2×2 matrix $A^{1/2}$ such that $A^{1/2} \cdot A^{1/2} = A$. From the symmetry pattern of A, we can see that $A^{1/2}$ will be in the form of $\begin{bmatrix} a & b \\ b & a \end{bmatrix}$, where a and b can be found by solving the following system of equations.

$$\begin{bmatrix} a & b \\ b & a \end{bmatrix} \cdot \begin{bmatrix} a & b \\ b & a \end{bmatrix} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

$$\begin{bmatrix} a^2 + b^2 & 2ab \\ 2ab & a^2 + b^2 \end{bmatrix} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}.$$

Thus, we have $a^2+b^2=1$ and $2ab=\rho$. After solving the system of equations for a and b, we obtain

$$a = \frac{\sqrt{1+\rho} + \sqrt{1-\rho}}{2} \text{ and } b = \frac{\sqrt{1+\rho} - \sqrt{1-\rho}}{2}.$$

As a result, if $V = \frac{n-2}{n} \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$, then $V^{1/2}$ will be given as

$$\frac{\sqrt{n-2}}{2\sqrt{n}} \begin{bmatrix} \sqrt{1+\rho} + \sqrt{1-\rho} & \sqrt{1+\rho} - \sqrt{1-\rho} \\ \sqrt{1+\rho} - \sqrt{1+\rho} & \sqrt{1+\rho} + \sqrt{1+\rho} \end{bmatrix}.$$

It is the expression given in (4.4).

Reference

Beasley, J.D. & Springer, S.G. (1977) Algorithm AS 111 : The percentage points of the Normal Distributions. Applied Statistics, 26, p.118-121.

Browne, M.W. (1982) Covariance Structure. In D. M. Hawkins (Ed.), Topics in Applied Multivariate Analysis (p.72-141). London : Cambridge University Press.

Cooper, B.E. (1968) Algorithm AS 3 : The integral of Student's t-distribution. Applied Statistics, 17, p.189-190.

Cox, N.R. (1974) Estimation of the correlation between a continuous and a discrete variable. Biometrics, 30, 171-178.

Donnelly, T.G. (1973) Algorithm 462 : Bivariate Normal Distribution, Communication of the ACM, 16, p.638.

Donnett, C.W. & Sobel, M.(1954) A Bivariate Generalization of Student's t-distribution, with Tables for Certain Special Cases, Biometrika, 41, p153-169.

Hannan, J. F. & Tate, R. F. (1965) Estimation of the parameters for a multivariate normal distribution when one variable is dichotomized. Biometrika, 52, 664-668.

Hill G.W. (1970) Algorithm 396 : Student's t-Quantiles, Communications of the ACM, 13, p.619-620.

Hill I.D. (1966) Algorithm AS 66 : The Normal Integral by I.D. Hill, Applied Statistics Algorithm, P.126-129.

Johnson N.L. & Kotz, S. (1972) Distributions in Statistics : Continuous Multivariate Distributions, New York : Wiley.

Joreskog, K.G. & Sorbom, D.G. (1982) LISREL User's Guide. Chicago : International Eductional Services.

Kennedy, W.J. Jr. & Gentle, J.E. (1980) Statistical Computing, New York : M. Dekker.

Lam M.L. (1987) Estimation of Polychoric Correlation with Non-normal Latent Variables. Master Thesis, Division of Statistics, The Chinese University of Hong Kong.

Lazarsfeld, P.F. (1959) Latent Structure Analysis, In S. Koch (Ed.) Psychology : A study of Science, Vol.3, New York : McGraw-Hill.

Lee, S.Y. & Poon, W.Y. (1986) Maximum Likelihood estimation of polyserial correlation. Psychometrika, 51, 113-121.

Marsaglia, G. (1961) Expressing a Random Variable in terms of Uniform Random Variables, Annals of Mathematical Statistics, 32, p.894-898.

Marsaglia, G. (1962) Random Variables and Computers, Information Theory Statistical Decision Functions Random Processes : Transactions of the Third Prague Conference, edited by J. Kozesnik, Czechoslovak Academy of Sciences, Prague, p.499-510.

Muirhead, R.J. (1982) Aspects of Multivariate Statistical Theory. New York : Wiley.

Olsson, U., Drasgow, F., & Dorans, N.J. (1982) The polyserial correlation coefficient. Psychometrika, 47, 337-347.

Poon, W.Y. & Lee, S.Y. (1987) Maximum Likelihood Estimation of Multivariate Polyserial and Polychoric Correlation Coefficients, Psychometrika, (1987). Vol. 52, No. 3, p.409-430.

Tate, R.F. (1955) The Theory of correlation between two continuous variable when one is dichotomized. Biometrika, 42, 205-216.